

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 1, 2003, 07:30:31 ; Search time 2775 Seconds  
(without alignments)  
10045.850 Million cell updates/sec

Title: US-10-049-742-22  
Perfect score: 1147  
Sequence: 1 gccttggtcaagcagaata.....gccttcctgagtagctggat 1147

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	895.2	78.0	942	13	BUS38214
2	796.8	69.5	2300	11	AK017830
3	796.8	69.5	2712	11	AK045445
4	795.2	69.3	2639	11	AK032681

5	795.2	69.3	4231	11	AK082579
6	792.6	69.1	918	13	BUS38214
7	787.4	68.6	3022	11	AK045345
8	786.4	68.6	4053	11	AK083878
9	764.6	66.7	957	12	BM802135
10	759.4	66.2	878	12	BG763500
11	744.2	64.9	955	13	BUS38351
12	706.4	61.6	785	12	B1333110
13	692.8	60.4	904	13	BQ691177
14	672.2	58.6	995	12	BM452174
15	644.2	56.2	723	10	BE619407
16	639	55.7	906	10	BG676002
17	622	54.2	1046	12	BM802913
18	605	52.7	605	13	BX090711
19	603.6	52.6	723	10	BE779250
20	601.8	52.5	940	10	BG328197
21	580	50.6	874	10	BF679562
22	573.8	50.0	927	12	BM424147
23	566	49.3	588	9	AA460131
24	565	49.3	1046	12	BQ58270
25	562.2	49.0	912	10	BG751089
26	559	48.7	737	14	CB324690
27	550	48.0	618	12	BM011379
28	548.6	47.8	805	10	BE536083
29	546.8	47.7	864	10	BG252385
30	541.4	47.2	679	9	AW146163
31	533.4	46.5	670	13	BU708368
32	530.4	46.2	667	14	BY760298
33	524.2	45.7	756	12	BG868350
34	520.4	45.4	900	10	BG298053
35	494	43.1	776	12	BG821631
36	490.4	42.8	651	10	BB465186
37	486.2	42.4	871	10	BF139916
38	478	41.7	560	9	AV602887
39	467	40.7	469	10	BG504969
40	465.2	40.6	527	12	B1002288
41	458.8	40.0	598	12	BM208252
42	436.6	38.1	448	12	B1002295
43	424.6	37.0	781	14	CB520550
44	419.4	36.6	505	10	BF819630
45	405	35.3	968	13	BUS38931

ALIGNMENTS

RESULT 1  
BUS38214  
LOCUS BUS38214 942 bp mRNA linear EST 13-SEP-2002  
DEFINITION AGENCOURT 10186479 NIH MGC 107 Homo sapiens cDNA clone IMAGE:6568738 5', mRNA sequence.  
ACCESSION BUS38214  
VERSION BUS38214.1 GI:22848655  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 942)  
AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1CM2754 row: 0 column: 10  
High quality sequence stop: 680.

```

FEATURES
  source      Location/Qualifiers
1. .942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6568738"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      236 a      235 c      244 g      222 t      5 others
ORIGIN
Query Match      78.0%; Score 895.2; DB 13; Length 942;
Best Local Similarity 97.4%; Pred.No. 2.7e-34;
Matches 917; Conservative 0; Mismatches 23; Indels 1; Gaps 1;

QY 182 GCCTATAGACAGCTGGCAGTGATGTTTCATCTGCACAAAATCATCATCCCGGGCTGAG 241
Db 1 GCCTATAGACAGCTGGCAGTGATGTTTCATCTGCACAAAATCATCATCCCGGGCTGAG 60

QY 242 GAGGCCTTCAAGGTTTTCGAGCAGCTTGGACATTTGACGAATGCTGAAAGCGAAAG 301
Db 61 GAGGCCTTCAAGGTTTTCGAGCAGCTTGGACATTTGACGAATGCTGAAAGCGAAAG 120

QY 302 GAGTATGAGTGAACGAAATGGCAGAGATGAGTGCAGCCGGTCAGTAAATGAGTTTCTG 361
Db 121 GAGTATGAGTGAACGAAATGGCAGAGATGAGTGCAGCCGGTCAGTAAATGAGTTTCTG 180

QY 362 TCCAGCTGCAAGTAGCTCAAGGAGGCAATGAACTACTATGATGTGACCGATGCCAA 421
Db 181 TCCAGCTGCAAGTAGCTCAAGGAGGCAATGAACTACTATGATGTGACCGATGCCAA 240

QY 422 GGAAAGCATAGGAGTTTGAATGGACCGGGAACTTAAGAGTGCCAGATCTGTGCTGAG 481
Db 241 GGAAGCATAGGAGTTTGAATGGACCGGGAACTTAAGAGTGCCAGATCTGTGCTGAG 300

QY 482 TGTATAGGCTGATCTCTGCTGAGGAAGAGAGACTTTTGGGCAGAGTCAAGCATGTTGGC 541
Db 301 TGTATAGGCTGATCTCTGCTGAGGAAGAGAGACTTTTGGGCAGAGTCAAGCATGTTGGC 360

QY 542 CTCAGATCACTACTTTTGCACGTATGATGGAAGGTCTATGACATCACAGATGGGCT 601
Db 361 CTCAGATCACTACTTTTGCACGTATGATGGAAGGTCTATGACATCACAGATGGGCT 420

QY 602 GGATCCGCGGTGATAGGTATCTCCCGAGATACCCACAGATCCCGCTATCACATCTCATTT 661
Db 421 GGATCCGCGGTGATAGGTATCTCCCGAGATACCCACAGATCCCGCTATCACATCTCATTT 480

QY 662 GGTTCCTGGATTCACAGGACCAAGGGCGGCAGAGAGCCACCCAGATGCCCTCTCTGCT 721
Db 481 GGTTCCTGGATTCACAGGACCAAGGGCGGCAGAGAGCCACCCAGATGCCCTCTCTGCT 540

QY 722 GATCTTCAGGATTTCTTGATGCGATCTTTCAAGTACCCCGAGGCGAGTCCCAATGGG 781
Db 541 GATCTTCAGGATTTCTTGATGCGATCTTTCAAGTACCCCGAGGCGAGTCCCAATGGG 600

QY 782 AACTTCTTTGAGCTCCTCAGCTCCCTGCGAGCGGCTGCGAGCTCTTAAGCCCAACAGC 841
Db 601 AACTTCTTTGAGCTCCTCAGCTCCCTGCGAGCGGCTGCGAGCTCTTAAGCCCAACAGC 660

QY 842 ACAGTACCCAGGGAGAGACCAAACTTAAGCGCGCGAAGAAAGTGAAGAGGCGCTTCCAA 901
Db 661 ACAGTACCCAGGGAGAGACCAAACTTAAGCGCGCGAAGAAAGTGAAGAGGCGCTTCCA 720

QY 902 CGTTGATGCCCTTCTCTTTTCCTCAAAATCAATGTGAGGGAGTCAAAAGGGCTGTAGCACA 961

```

```

Db 721 CGTTGATGCCCTTCTCTTTTCCTCAAAATCAATGTGAGGGTCAAAAGGGCTGTAGCACA 780
QY 962 GGATGGAGTTTGATTTATCCCTCCTCCCAACACCTAGGAACCTGAATCTTTT-CTTTT 1020
Db 781 GGATGGAGTTTGATTTATCCCTCCTCCCAANANCTAAGAACCTGAATCTTTTCTTTT 840
QY 1021 TATTTTGTGAGATGAGTCTTGTCTCTGTTGCCAGTGGAGTGCAGTGGTGTGATCTCAG 1080
Db 841 TATTTTGTGAGATGAGTCTTGTCTCTGTTGCCANCTGGATGTCANTGGTGTGATCTTCA 900
QY 1081 CTTACTGCAACCTCTGTCTCCCGGGTTCAAGCAATTTCTCC 1121
Db 901 GCTACTGCCACCTGTCTCCCGGGTTCCAGCAATTTCTCC 941

RESULT 2
AK017830      2300 bp      mRNA      linear      HTC 05-DEC-2002
LOCUS      Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
DEFINITION      enriched library, clone:5730551F12 product:DOPAMINE RECEPTOR
INTERACTING PROTEIN homolog [Rattus norvegicus], full insert
sequence.
ACCESSION      AK017830
VERSION      AK017830.1 GI:12857283
KEYWORDS      HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1
AUTHORS      Carninci, P. and Hayashizaki, Y.
TITLE      High-efficiency full-length cDNA cloning
JOURNAL      Meth. Enzymol. 303, 19-44 (1999)
MEDLINE      99279253
PUBMED      10349636
REFERENCE      2
AUTHORS      Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE      Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL      Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE      20499374
PUBMED      11042159
REFERENCE      3
AUTHORS      Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE      RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL      Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE      20530913
PUBMED      11076861
REFERENCE      4
AUTHORS      Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yanakita, I.,
Saito, K., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staib, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Nordone, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Sasaki, H.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

```

Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

AUTHORS  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (baes 1 to 2300)

AUTHORS  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukumoto, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL

Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATCCAGTAAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

source

Location/Qualifiers  
1..2300  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:5730551F12"  
/db\_xref="MGI:1907484"  
/db\_xref="caxon:10090"  
/clone="5730551F12"  
/tissue type="whole body"  
/clone lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="8 days embryo"  
97..2208

CDS

/note="unnamed protein product; DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SPR|Q925G7, evidence: FASTA, 92.7%ID, 100%length, match=2109)  
putative"  
/codon\_start=1  
/protein\_id="BAB30962.1"  
/db\_xref="GI:12857284"  
/db\_xref="MGI:1921580"

/translation="MAQKHGERRRLCGAHRSGTSLSTSGSSVDPEILSFGSLRSDAE  
TAPNCRCLKSHSGPKYTPPNPAHSDPSHGPPRGPGYGGVDESETSEBSGV  
DOELSENETGYQEDGSPFSLIPACNQCSPGVEGTYYSEGGSSSSSLCHCTSP  
ALGDELEEEYDDDEELKFFSDFSRVSSGKPLSRKQRLFLIEDKSDSRSPKPA  
PGRHLARKSQTDKRGGLGWBEELCOLGOAGFWMLIELLVGEYVETCGHLIYA  
CRKLKSGDLDFRVWVGWARRLGGWARMFQFSQFFCVGLLIRIVVGAFLIL  
ALALFLGCLQGLWRFVSLGNRLGWRDKTAMLFSLGSPALHCLTLKLDSPWQOLV  
RLIQWQELPKVKQRTKQGNAPVSNRGYCOPEEVTPLLTLMAGVPDELPFVILG  
VEATSDIELKAYRQLAVMVHPDKNHPHPRABEAFKILRAAMDIVSNPRKREYMKR  
MAENLSRSVNEFSLKLDLDDKAMNTWMSRCQKHRRFEMDRPKSYCAECNRL  
HPAEEGDFWASSMLGLKITYPALMDGVYDITWAGQORVIGSPDTHRPVPHISFGS  
RVPGTSGQRATPEPPADLDQFLSRIFQVPPGPMNSNGFFAAPHGPGTGTSTSRNS  
SVPGKEAKPKRRKKVRRPQR"

BASE COUNT 557 a 560 c 675 g 508 t  
ORIGIN

Query Match 69.5%; Score 796.8; DB 11; Length 2300;  
Best Local Similarity 89.5%; Pred. No. 5.8e-30;  
Matches 982; Conservative 0; Mismatches 97; Indels 7; Gaps 2;  
  
QY 1 GCCTTGGGTCAAGCAGATAATTTAATAGGCGGGGAATGCACCTGTAGCTAGTGGGGCGCTA 60  
DB 1302 GCCCTGGGTCAAAACAGAGGACTAAGAAACAGGGCAATGCACCTGTAGCTAGCGGGCGATA 1361  
QY 61 CTGCCAGCCTGAAGAGGAAGTGGCTCTTGACCATGGCTGGGTTCCTTGAGGATGA 120  
DB 1362 CTGCCAGCCTGAAGAGGAAGTGCACCGACTCTTGACCATGGCTGGGTTCCTTGAGGATGA 1421  
QY 121 GCTAAACCTTTCCTATGCTTGGGGTTCAGCCACAGCATCAGATGTTGAACGTGAAGAA 180  
DB 1422 ACTAAACCTTTCCTATGCTTGGGGTTCAGCTACAGCATCAGCATCGACACTGAAGAA 1481  
QY 181 GGCCTATAGACAGCTGGCAGTGCATGGTTCATCCTGACAAAATCATCATCCCGGGCGCTGA 240  
DB 1482 GGCCTATAGGAGCTAGCAGTAATGGTCCATCCTGATAAAATCACCATCCCGGGCGCTGA 1541  
QY 241 GGAGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATTTGACCAATGCTGAAAAGCGCAA 300  
DB 1542 GGAGCGCTTCAAAATTTTCGCGGCGAGCTGGGACATTTGACCAACCCAGAGAGCGGAA 1601  
QY 301 GGAGTATGAGATGAAACGAATGGCAGAGATGAGCTGAGCCGGTCAGTAAATGAGTTTCT 360  
DB 1602 GGAATATGAGATGAAACGAATGGCAGAGATGAGCTGAGCCGGTCAGTAAATGAGTTTCT 1661  
QY 361 GTCCAAGCTGCAAGATGACCTCAAGGAGCAATGAATGATGATGTGTAGCGGATGCCA 420  
DB 1662 GTCGAAATACCAAGATGACCTCAAGGAGCAATGAACACGATGATGTGACAGAGATGCCA 1721  
QY 421 AGGAAAGCATAGGAGTTTGAATGCGACCGGAACTTAAGAGTGCACAGATCTGTGCTGA 480  
DB 1722 AGGAAAGCATAGGAGTTTGAATGCGACCGGAACTTAAGAGTGCACAGATCTGTGCTGA 1781  
QY 481 GTGTAATAGGCTGCATCCTGCTGAGAAAGGAGACTTTTGGGCGAGAGTCAAGCATGTTGGG 540  
DB 1782 GTGTAACAGGCTGCATCCTGCTGAGAAAGGAGACTTTTGGGCGAGAGTCCAGCATGCTGGG 1841  
QY 541 CCTCAAGATCACTTACTTTGCACTGATGGATGGAAGAGTGTATGACATCACAGAGTGGGC 600  
DB 1842 CCTCAAGATCACTTACTTTGCGCTGATGGATGGAAGAGTGTATGACATCACAGAGTGGGC 1901  
QY 601 TGGATGCCAGCGTGTAGTATCTCCCCAGATACCCACAGAGTCCCTTATCATCTCATTT 660  
DB 1902 TGGATGCCAGCGTGTGGGTATCTCCCCAGATCTCCACAGAGTTCCTTACCATCTCATTT 1961  
QY 661 TGGTTCTCGGATTCAGGACACAGAGGCGGCGAGAGAGCCACCCAGAGTCCCTCTCTGC 720  
DB 1962 TGGTTCTCGGATTCAGGACACAGTGGCGGCGAGAGGCGCACTCCAGAGTCCCTCTCTGC 2021  
QY 721 TGATCTTCAGGATTTCTTTAGTTCGATCTTTCAAGTACCCCGGCGAGATGCCCAATGG 780  
DB 2022 TGACTGCGAGATTTCTTTAGCGGATCTTTCAAGTACTCTCGGGGCGGATGTCCTCAATGG 2081  
QY 781 GAACTTCTTTGAGCTCTCTAGCCTGCCCCCTGGAGCCGCTGAGCCTCTAAAGCCCAACAG 840



MANELSRSVNELSKLQDQLKEAMNTMCSRQCKXHRFEMDREPKSARYCAECNRL  
HPAEEDGFWAESSMLGIKITYFALMDCKVYDITEWAGCORVG:SPOTHRVPYHSIGS  
RVPGTSGORAPATPESPADLQDSRIIFQVPPGPMGNGNFFAAPHQPGTITSIRPNS  
SVBKGEAKPKRKKVRPFOR"

664 a 663 c 778 g 607 t

BASE COUNT  
ORIGIN

Query Match.	69.5%;	Score 796.8;	DB 11;	Length 2712;
Best Local Similarity	89.5%;	Pred. No. 5.2e-30;		
Matches 882;	Conservative 0;	Mismatches 97;	Indels 7;	Gaps 2;
QY	1	GCCTTGGGTCACGACGAATATTATATGGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA	60	
Db	1382	GCCTTGGGTCACGACGAATATTATATGGCAGGGGAATGCACCTGTAGCTAGTGGCGGATA	1441	
QY	61	CTGCCAGCCTGAAGAGGAAGTGGCTGCACCTCTTGACCATGGCTGGGTTCTCTGAGGATGA	120	
Db	1442	CTGCCAGCCTGAAGAGGAAGTGGCTGCACCTCTTGACCATGGCTGGGTTCTCTGAGATGA	1501	
QY	121	GCTAAACCCCTTTCCATGTATCTGGGGGTTGAGGCCACAGCATTCAGATGTTGAACCTGAAGAA	180	
Db	1502	ACTAAACCCCTTTTCATGTGTGGGGGTTGAGGTTACAGCATCCGACACTGAACCTAAAGAA	1561	
QY	181	GGCCTATAGACAGCTGGCAGTGATGGTTTCATCTCTGACAAAATCATCATCCCGGGCTGA	240	
Db	1562	GGCCTATAGCAGCTAGCAGTAATGGTCCATCTCTGATAAAAATCACCATCCCGGGCTGA	1621	
QY	241	GGAGGCCCTTCAAGGTTTTTGGCAGCAGCTTGGGACATTTGTCCAGCAATGCTGAAAAGCGAAA	300	
Db	1622	GGAGGCCCTTCAAAAATTTTGGCGGCAGCTTGGGACATTTGTCCAGCAATGCTGAAAAGCGGAA	1681	
QY	301	GGAGTATGAGATGAACCAATGGCAGAGATGAGCTGAGCGGTGAGTAAATCAGTTTTCT	360	
Db	1682	GGAATATGAGATGAACCGGATGGCAGAGATGAGCTCAGCGGTGAGTAAATCAGTTTTCT	1741	
QY	361	GTCCAAGCTGCAGATGACCTCAAGGAGGCAATGAATACTATGATGTGAGCCGATGCCA	420	
Db	1742	GTCCAAACTCAAGATGACCTCAAGGAGGCAATGAACCATGATGTGACGACGATGCCA	1801	
QY	421	AGGAAAGCATAGAGAGGTTTCAAAATGGACCGGGAACCTTAAGAGTGCACAGATACTTGCTGA	480	
Db	1802	AGGAAAGCATAGAGGTTTCAAAATGGACCGGGAACCTTAAGAGTGCACAGATACTTGCTGA	1861	
QY	481	GTGTAATAGGCTGCATCTCTCTGAGGAAGAGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	540	
Db	1862	GTGTAACAGGCTGCATCTCTCTGAGGAAGAGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	1921	
QY	541	CCTCAAGATCACCTACTTTTGCATGTATGGATGGAAAGGTGTATGACATCACAGTGGGC	600	
Db	1922	CCTCAAGATCACCTACTTTTGCATGTATGGATGGAAAGGTGTATGACATCACAGTGGGC	1981	
QY	601	TGATGCCAGGTGTAGGTATCTCCACAGATACCCACAGAGTCCCTTATCACATCTCATT	660	
Db	1982	TGATGCCAGGTGTAGGTATCTCCACAGATATCTCCACAGATATCTTATCACATCTCATT	2041	
QY	661	TGTTTCTCGGATCTCCAGGCCACAGAGGGCGGAGAGCCACCCACAGATGCCCTCTCTGC	720	
Db	2042	TGTTTCTCGGATCTCCAGGCCACAGAGGGCGGAGAGCCACCCACAGATGCCCTCTCTGC	2101	
QY	721	TGATCTTCAGGATTTCTTGAGTCGGATCTTTTCAAGTACCCCGAGGCGAGATGCCAATGG	780	
Db	2102	TGACCTTCAGGATTTCTTGAGTCGGATCTTTTCAAGTACCTCCGGGCGGATGTCCAAATGG	2161	
QY	781	GAACTTTCTTGAGCTCTCTCAGCCTGCCCTGGAGCGGCTGCAGCTCTTAAGCCCAACAG	840	
Db	2162	GAACTTTCTTGCGCACCTCACCTGCCCTGGGACCATTTGACCTCTTAGGCCCAACAG	2221	
QY	841	CACAGTACCCAGGGAGAGCCAAAACCTAAGCGGCGGAAGAAAGTGGAGGGCCCTTCCA	900	
Db	2222	TTTCAATACCCAGGGAGAGCCAAAACCTAAGCGGCGGAAGAAAGTGGAGGGCCCTTCCA	2281	
QY	901	AGTTGATGCCCTTTCTCT---TTCCTCAAAATCAATGTCAGGGAGTCAAAAGGGCTGT--	955	

Db	2282	ACGATGACACCCCTTCTCTTCTCTCTCAATGTCAGGGAGTCAAAAGGGCTGTGT	2341
Qy	956	--AGCAGGATGGAGTTTGATTAT	979
Db	2342	ACAGCAGGATGGATTTGATTAT	2367
RESULT 4			
LOCUS	AK032681	2639 bp	mRNA linear HTC 05-DEC-2002
DEFINITION	Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:6530416L18 product:DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus], full insert sequence.		
ACCESSION	AK032681		
VERSION	AK032681.1	GI:26082970	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Math. Enzymol. 303, 19-44 (1999)		
MEDLINE	95273253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		

AUTHORS	TITLE
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komoro, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakata, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schmitt, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection
	Nature 409 (6821), 685-690 (2001)
	21095660
	11217851
	REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2639)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imozumi, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichou, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers

FEATURES

source 1..2639 /organism="Mus musculus" /mol\_type="mRNA" /strain="C57BL/6J" /db\_xref="FANTOM DB:6530416L18" /db\_xref="taxon:10090" /clone="6530416L18" /tissue\_type="cerebellum" /clone\_libs="RIKEN full-length enriched mouse cDNA library" /dev\_stages="10 days neonate" misc\_feature 64..2176 /note="DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SPTR|Q925G7, evidence: FASTA, 92.%ID, 100%length, match=2109) putative"

BASE COUNT 666 a 634 c 738 g 601 t

ORIGIN

Query Match 69.3%; Score 795.2; DB 11; Length 2639; Best Local Similarity 89.4%; Pred. No. 6.3e-30; Matches 881; Conservative 0; Mismatches 98; Indels 7; Gaps 2;

QY 1 GCCTTGGTCAAGCAGATATTAATAGCAGGGGATGCACCTCTAGTGTGGCGCTGA 60

DB 1270 GCCCTGGTCAACAGAGGATTAAGAAACAGGGCAATGCACCTGTAGCTAGCGGCGGATA 1329

QY 61 CTGCCAGCCTGAAGAGGAAGTGGCTCGACTCTTGACCATGCTGGGGTTCCTGAGGATGA 120

DB 1330 CTGCCAGCCTGAAGAGGAAGTGAACCGACTCTTGACCATGCTGGGGTTCCTGAGGATGA 1389

QY 121 GCTAAACCTTTTCATGTACTGGGGTTAGGCCACAGCATCAGATGTTGAACCTGAAGAA 180

DB 1390 ACTAAACCTTTTCATGTCGTGGGGTTGAAGCTACAGCATCCGACACTGAACCTAAAGAA 1449

QY 181 GGCCTATAGCAGCTGGCAGTGATGTTTCATCCTGACAAAATCATATCCCGGGCTGA 240

DB 1450 GGCCTATAGCAGCTAGCAGTAATGGTCCATCTCTGATAAAAAATCACCATCCCGGGCTGA 1509

QY 241 GGAGGCCCTTCAAGGTTTTCGAGCAGCTTGGGACATTTGTGACAAATGCTGAAAGCGGAA 300

DB 1510 GGAGGCCCTTCAAAATTTTTCGGGCGAGCTTGGGACATTTGTGACAAACCCAGAGAGCGGAA 1569

QY 301 GGAGTATGAGATGAACGAAATGGCAGAGAAATGAGTGTAGCCGGTTCAGTAAATGAGTTTCT 360

DB 1570 GGAATATGAGATGAACGAGTGGCAGAGAAATGAGTGTAGCCGGTTCAGTGAATGAGTTTCT 1629

QY 361 GTCCAAAGCTGCAAGATGACCTCAAGAGGCAATGAATGATGATGTTAGCGGATGCCA 420

DB 1630 GTCCGAAATCTACAAGATGACCTCAAGAGGCAATGAACACGATGATGTGCGAGATGCCA 1689

QY 421 AGGAAGCATAGGAGGTTTGAATGGACCGGAACTTAAGAGTGCAGATATCTGCTCTGA 480

DB 1690 AGGAAGCATAGGAGGTTTGAATGGACCGGAACTTAAGAGTGCAGATATCTGCTCTGA 1749

QY 481 GTGTAATAGGCTGCATCTCTGCTGAGGAAAGGACACTTTTGGGCGAGAGTCAAGCATGTTGGG 540

DB 1750 GTGTAACAGGCTGCATCTCTGCTGAGGAAAGGACACTTTTGGGCGAGAGTCAAGCATGCTGGG 1809

QY 541 COTCAAGATCACTACTTTGCACTGATGGATGGAAAGGTGTATGACATCACAGAGTGGGC 600

DB 1810 COTCAAGATCACTACTTTGCGCTGATGGATGGAAAGGTGTATGACATCACAGAGTGGGC 1869

QY 601 TGGATGCCAGGCTGTAGTATCTCCCGAGATACCCACAGAGTCCCTTATCAGATCTCATTT 660

DB 1870 TGGATGCCAGGCTGTGGGTATCTCCCGAGATACCTCACAGAGTTCCTTACACATCTCATTT 1929

QY 661 TGGTTCTCGGATTCAGGACCAAGGCGGCGAGAGGCCACCCAGATGCCCTCCCTGCTGC 720

DB 1930 TGGTTCTCGGTACCCGCGACCCAGTGGCGGCGAGAGGCCACTCCAGAGTCCCTCCCTGCTGC 1989

QY 721 TGATCTTCAGGATTTCTTTGAGTCGGATCTTTCAAGTACCCCCAGGCGAGATGCCCAATGG 780

DB 1990 TGACCTGCAGAGATTTCTTTGAGCCGGATCTTTCAAGTACTCCGGGCGAGTGTCCATGG 2049

QY 781 GAACCTTTTTCAGCTCTCAGCTTGCCTTGGAGCCCTCGAGCTCTTAAGCCCCAACAG 840

DB 2050 GAACCTTTTTCGCGCACCTCACCCTGGCCTGGGACCACTTCGACCTCTAGGCCAACAG 2109

QY 841 CACAGTACCCAGGAGGAGAACCAACCTTAAGCGGCGGAGAAAGTGAAGAGGCCCTTCCA 900

DB 2110 TTCAAGTACCCAGGAGGAGAACCAACCTTAAGCGGCGGAGAAAGTGAAGAGGCCCTTCCA 2169

QY 901 ACGTTGATGCCCTTCTCTCT---TTCTCTCAATCAATGTCAAGGAGTCAAAAGGGCTGT-- 955

DB 2170 ACGATGACACCCCTTCTCTCTCTCTCTCAATCAATGTCAAGGAGTCAAAAGGGCTGTGT 2229

QY 956 --AGCACAGGATGGAGTTTGATTTAT 979

DB 2230 ACAGCACAGGATGGAGTTTGATTTAT 2255

RESULT 5

AK082579

LOCUS AK082579 4231 bp mRNA linear HTC 05-DEC-2002

DEFINITION Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched library, clone: C230066C20 product: DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus], full insert sequence.

ACCESSION AK082579

VERSION AK082579.1 GI:26349792

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 MEDLINE 10349636 PUBMED

REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res.	10 (10), 1617-1630 (2000)
MEDLINE	20499374	
PUBMED	11042159	
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS		RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res.	10 (11), 1757-1771 (2000)
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Horrmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.
AUTHORS		Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature	409 (6821), 685-690 (2001)
MEDLINE	21085660	
PUBMED	11217851	
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS		Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature	420, 563-573 (2002)
MEDLINE		6 (bases 1 to 4231)
PUBMED		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, N., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
AUTHORS		Direct Submission
JOURNAL	Submitted (16-APR-2002)	Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
REFERENCE		cDNA library was prepared and sequenced in Mouse Genome
COMMENT		

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

1. .4231

for organism="Mus musculus"

/mol\_type="mRNA"

/strain="CS7BL/6J"

/db\_xref="FANTOM\_DB:C230066C20"

/db\_xref="taxon:10090"

/clone="C230066C20"

/tissue\_type="cerebellum"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="0 day neonate"

84. .2195

/note="unnamed protein product; DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus]"

(SPTS)Q925G7, evidence: FASTV, 92.7%ID, 100%length, match=2109)

putative"

/codon\_start=1

/protein\_id="BAC38536.1"

/db\_xref="GI:26349793"

/translation="MAQKHGPRRLCGAHRSGTSLSTSGSVDPILSFSGLRDSAE TAPNTRCLKHSHPKTYTPNPAPHSWDPHPGPPGPGVDESETSESEGV DQLSRENETGYDEDSPLFSIPACNCGSPGVPEGTYSSEGGSSSLCHCTSP ALGDELEEEYDDPELFPDSFRVSKKPLGRQKHRLIKEDVDSTGRPEKA PGRHELARKSOTDKRGLGMLGVELCOLGOAGFWMLJELLVLVEYVETCGLLIYA CRKLKGLDLDFRVVGVWARRLGGWARMFQSLQSFVGVGLLIRLRVCAFILLL ALALFLGCLQLGWRFSVGLNRLGWDKTAWLFSWGLSPALHHLCLLLKDSRPWQQLV RLIOGWDELPMVKQRTKKQGNAPVAGRYCQPEBEVRLTMAGVPDELNPFVLG VEATASDELKAYKQLAWPHDPKNNHPRAEAEFKILAAWDIVSNPERRKEYEMKR MAENLSRVNEFLSLQDLLEAMNTWMSRCQGRHREFMREDRPKSAGYCAECNRL HPAEESDPAWSSMLGLKITITPALMDNKYDITWAGCQGVGISPDRVPHISFGS RVPFGSRQATPESPADLQDLFSRIFQVPPGPMNSGNFFAHPGPGTSTSTSRNS SVPKGEAKPKRKKVRPFQR"

BASE COUNT

1092 a 948 c 1059 g 1132 t

ORIGIN

Query Match 69.3%; Score 795.2; DB 11; Length 4231;

Best Local Similarity 89.4%; Pred. No. 4.6e-30;

Matches 881; Conservative 0; Mismatches 98; Indels 7; Gaps 2;

Qy 1 GCCTTGGGTCAAGCAGAGATATTAATAGGCAGGCGGATGCACCTGTAGCTAGTGGCGCTA 60

Db 1289 GCCTTGGGTCAAGCAGAGACTAAGAAACAGGCGGCAATGCACCTGTAGCTAGCGGCGATA 1348

Qy 61 CTGCCAGCCTCAAGAGGAAGTGGCTCGACTCTTGACCATGGCTGGGGTTCCTGAGGATGA 120

Db 1349 CTGCCAGCCTCAAGAGGAAGTGCACCGACTCTTGACCATGGCTGGGGTTCCTGAGGATGA 1408

Qy 121 GCTAAACCCCTTCCATGTAATGCGGGTGTAGGCCACAGCATCAGATGTTTGAATGAAGAA 180

Db 1409 ACTAAACCCCTTTCATGTGCTGGGGTGTGAAGCTACAGCATCCGACACTGAACCTAAAGAA 1468

Qy 181 GGCCTATAGACAGCTGGCAGTGCATGGTTCCTCGACAAAATCATCATCCCGGGCTGA 240

Db 1469 GGCCTATAGCAGCTAGCAGTAATGGTTCCTCGATAAAAATCACCATCCCGGGCTGA 1528

Qy 241 GGAGCCCTCAAGGTTTTCGCGAGCAGCTGGGACATTTGTCAGCAATGCTGAAAGCGGAA 300

Db 1529 GGAGCCCTCAAAATTTTCGGGGCAGCTGGGACATTTGTCAGCAACCCAGAGAGCGGAA 1588

Qy 301 GGAGTATGAGATGAACCAATGGCAGAGAATCAGCTGAGCCGCTCAGTAAATGAGTTTCT 360

Db 1589 GGAATATGAGATGAACCGATGGCAGAGAATGAGCTCAGCCGCTCAGTAAATGAGTTTCT 1648

Qy 361 GTCCAGCTGCAAGATGACCTCAAGGAGCAATGAATACTATGATGTGTAGCCGATGCCA 420







QY	1014	TTCTTTTATTTTTCAG-ATGAGCTTGTCTC-TGTTGCCAGCTGGAGT 1062	
	Db		
RESULT 7	AK045345	3022 bp mRNA linear HTC 05-DEC-2002	
	LOCUS	Mus musculus adult male corpora quadrigenima cDNA, RIKEN	
DEFINITION	AK045345	full-length enriched library, clone:B230104C23 product:DOPAMINE	
	RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus], full		
ACCESSION	AK045345	insert sequence.	
	VERSION	AK045345.1 GI:26337270	
KEYWORDS	HTC; CAP trapper.		
	SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS	HTC	High-efficiency full-length cDNA cloning	
	JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE	99279253		
	PUBMED	10349636	
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
	AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
	MEDLINE	20499374	
PUBMED	11042159		
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
	AUTHORS	Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,	
TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
JOURNAL	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,		
	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,		
MEDLINE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
	PUBMED	11076861	
REFERENCE	4	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,	
	AUTHORS	Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,	
TITLE	Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,		
	Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,		
JOURNAL	Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,		
	Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,		
MEDLINE	Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,		
	PUBMED	11076861	
REFERENCE	5	Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M.,	
	AUTHORS	Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Bono, H.,	
TITLE	Baldarelli, R., Barsh, G.D., Blake, J., Boffelli, D., Bojunga, N.,		
	Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,		
JOURNAL	Fleischer, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,		
	Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,		
MEDLINE	Marchionni, L., Mashima, J., Mazzarelli, J., Monbaerts, P., Nordone, P.,		
	PUBMED	11076861	
REFERENCE	6	Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,	
	AUTHORS	Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,	
TITLE	Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,		
	Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohtsuki, S.		
JOURNAL	and Hayashizaki, Y.		
	PUBMED	11076861	
REFERENCE	7	Functional annotation of a full-length mouse cDNA collection	
	AUTHORS	Nature 409 (6821), 685-690 (2001)	
MEDLINE	21085660		
	PUBMED	11217851	
REFERENCE	8	The FANTOM Consortium and the RIKEN Genome Exploration Research	
	AUTHORS	Group Phase I & II Team.	

TITLE	Analysis of the mouse transcriptome based on functional annotation		
	of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
	REFERENCE	6 (bases 1 to 3022)	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,		
	Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,		
TITLE	Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,		
	Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,		
JOURNAL	Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,		
	Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,		
MEDLINE	Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,		
	PUBMED	10349636	
REFERENCE	9	Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,	
	AUTHORS	Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,	
TITLE	Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,		
	Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,		
JOURNAL	Muramatsu, M. and Hayashizaki, Y.		
	PUBMED	11076861	
REFERENCE	10	Direct Submission	
	AUTHORS	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of	
TITLE	Physical and Chemical Research (RIKEN), Laboratory for Genome		
	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		
JOURNAL	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,		
	Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,		
MEDLINE	URL: http://genome.gsc.riken.go.jp/; Tel: 81-45-503-9222,		
	PUBMED	10349636	
COMMENT	CDNA library was prepared and sequenced in Mouse Genome		
	Encyclopedia Project of Genome Exploration Research Group in Riken		
FEATURES	Genomic Sciences Center and Genome Science Laboratory in RIKEN.		
	Division of Experimental Animal Research in Riken contributed to		
SOURCE	prepare mouse tissues.		
	Please visit our web site for further details.		
CDS	URL: http://genome.gsc.riken.go.jp/		
	URL: http://fantom.gsc.riken.go.jp/		
FEATURES	Location/Qualifiers		
	1..3022		
source	/organism="Mus musculus"		
	/mol_type="mRNA"		
CDS	/db_xref="FANTOM DB:B230104C23"		
	/db_xref="taxon:10090"		
FEATURES	/clone="B230104C23"		
	/sex="male"		
CDS	/tissue_type="corpora quadrigenima"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
FEATURES	/dev_stage="adult"		
	140..2176		
CDS	/note="unnamed protein product; DOPAMINE RECEPTOR		
	INTERACTING PROTEIN homolog [Rattus norvegicus]		
FEATURES	(SPT)Q925G7, evidence: FASTV, 92.7%ID, 100%length,		
	match=2109		
CDS	putative"		
	/codon_start=1		
FEATURES	/protein_id="BAC32320.1"		
	/db_xref="GI:26337271"		
CDS	/translation="MAQKHGPERRLCGAHRSGTSLTSGSSVDPEILSFGURDSAE		
	TAPNTRCLKHSQPKYQPNPAHSDPSRGPGRPGPGGYPDSESTSESGV		
FEATURES	DOESRENETGQDGSFUSIPACNCGSPGVEGTYSSEGGSSSLCHCTSP		
	ALGDELEEDYDEEPLKFPDSFVSSGKPLSRQKRLFLVDEVDSSRRFPFA		
CDS	PGHRLARKSQTDRGLGVEELQQLQAGFWMLIELLVGEYVETCGHIIYA		
	CRKLGSGLDLFRVWVGWARRLGMARMFQFSQFPCVVGLLIRLVVGAFLLL		
FEATURES	ALALFLGCLQLGWRFPSVGLGNRLGWRDXTAMLFSLGSPALHHCITLLKDSFPMQOLV		
	RLIQWGLFWKQRTKQGNAPVAGRYQPEEVEVRLTMAGUPEDELNPFVLG		
CDS	VEATSDTELKAYQLAVMHPDKNHPRAEAFKILRAWDIVSNPERKEYENKR		
	MAENLSVSNVEFSLQDDIKAEKNTMWCRCQGRFRFMDRFPKSAKCAEENRL		
FEATURES	HPAEGDFWAKSSMLGKITFYALMDGKVDITWAGQRCVIGISPTIRVPVHIFGS		
	RVPTGSGRQATPESPFPADLQDFLSIRFIQVFSGADVQVWELLCRTSPMPWHDHFL		
BASE COUNT	753 a 716 c 826 g 727 t		
	ORIGIN		
Query Match	68.6%; Score 787.4; DB 11; Length 3022;		
	Query Local Similarity 89.5%; Pred. No. 1.3e-29;		
Matches	883; Conservative 0; Mismatches 96; Indels 8; Gaps 3;		
	1 GCCTTGGTCAAGCAGAATATTAATAGCAGGGGAATGCACCTGTAGTGGCGCTA 60		



Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
FEATURES  
source  
1. 4053  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB: D130038P18"  
/db\_xref="taxon:10090"  
/clone="D130038P18"  
/tissue\_type="spinal ganglion"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
misc\_feature  
94..2205  
/notes="DOPAMINE RECEPTOR INTERACTING PROTEIN homolog [Rattus norvegicus] (SEPR|Q925G7, evidence: FASTV, 92.7%ID, 100%length, match=2109)  
putative"  
BASE COUNT 1060 a 903 c 1039 g 1051 t  
ORIGIN  
Query Match 68.6%; Score 786.4; DB 11; Length 4053;  
Best Local Similarity 89.5%; Pred. No. 1.2e-29;  
Matches 882; Conservative 0; Mismatches 96; Indels 8; Gaps 3;  
QY 1 GCCTGGGTCAACGAGAAATTAATAGGCGAGGGGAATGACCTGTAGCTAGTGGGGCTGA 60  
Db 1300 GCCTGGGTCAACGAGGAGCTAAGAAACAGGGCAATGCACCTGTAGCTAGCGGGCGATA 1359  
QY 61 CTGCCAGCCTGAGAGGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCCTCAGCATGA 120  
Db 1360 CTGCCAGCCTGAGAGGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCCTCAGCATGA 1419  
QY 121 GCTAAACCCCTTTCATGTACTGGGGGTTGAGGCGCACGACATCAGATGTTGAACCTGAAGAA 180  
Db 1420 ACTAAACCCCTTTCATGTACTGGGGGTTGAGGCGCACGACATCAGATGTTGAACCTGAAGAA 1479  
QY 181 GGCTATAGACAGCTGGCAGTGTGTTTCATCTCGACAAATAATCATCTCCCGGGCTGA 240  
Db 1480 GGCTATAGCAGCTAGCAGTGTGTTTCATCTCGACAAATAATCATCTCCCGGGCTGA 1539  
QY 241 GGAGGCTTCAAGTTTTCGGAGCAGCTTGGGACATGTCAGCAATGCTGAAAGCGAA 300  
Db 1540 GGAGGCTTCAAAATTTTCGGGACGCTTGGGACATGTCAGCAATGCTGAAAGCGGAA 1599  
QY 301 GGAGTATGAGATGAACAGGAATGCGCAGAGATGAGTGGCGGTTCAGTAAATGAGTTTCT 360  
Db 1600 GGAATATGAGATGAACAGGAATGCGCAGAGATGAGTGGCGGTTCAGTAAATGAGTTTCT 1659  
QY 361 GTCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATATGATGTAGCGCATGCCA 420  
Db 1660 GTCGAACTACAAGATGACCTCAAGAGGCAATGAACAGATGATGTGAGCGATGCCA 1719

QY 421 AGCAAGCATAGAGGCTTTGAAATGACCGGGAACCTAAGAGTGCAGATACTGTGCTGA 480  
Db 1720 AGCAAGCATAGAGGCTTTGAAATGACCGGGAACCTAAGAGTGCAGATACTGTGCTGA 1779  
QY 481 GTCTAATAGCTCATCTCCTGCTGAGGAAGGAGACTTTTGGGCGAGAGTCAAGACATGTTGGG 540  
Db 1780 GTCTAATAGCTCATCTCCTGCTGAGGAAGGAGACTTTTGGGCGAGAGTCAAGACATGTTGGG 1839  
QY 541 CTTCAAGATCACCTACTTTTGCACTGATGTAAGAAAGGTGTATGACATCAAGAGTGGGC 600  
Db 1840 CTTCAAGATCACCTACTTTTGCGCTGATGTAAGAAAGGTGTATGACATCAAGAGTGGGC 1899  
QY 601 TGGATGCCAGCGTGTAGGTATCTCCAGATACCCAGAGTCCCTTATCACATCTCTCAT 660  
Db 1900 TGGATGCCAGCGTGTGGGTATCTCCAGATACCTCAGAGTTCCTTACCACATCTCAT 1959  
QY 661 TGGTCTCGGATTCAGGACACAGAGGGCGGAGAGGCCACCCAGATGCCCTCTCTGC 720  
Db 1960 TGGTCTCGGATTCAGGACACAGAGTGGCGGAGAGGGCCACTCCAGA-GTCCCTCTCTGC 2018  
QY 721 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCAGGCGAGAGTGCCTAATGG 780  
Db 2019 TGACCTGCGAGGATTTCTTGAGTCGGATCTTTCAAGTACCTCCGGGCGGATGTTCCAATGG 2078  
QY 781 GAACCTTTTTCAGCTCCTCAGCTGCTCCCTGGAGCGGCTGCGAGCTCTTAAGCCCAACAG 840  
Db 2079 GAACCTTTTTCAGCTCCTCAGCTGCTCCCTGGAGCGGCTGCGAGCTCTTAAGCCCAACAG 2138  
QY 841 CACAGTACCAAGGAGAGAGCCAAACCTAAGCGCGGAGAAAGTGAAGAGGCGCTTCCA 900  
Db 2139 TTCAAGTACCAAGGAGAGAGCCAAACCTAAGCGCGGAGAAAGTGAAGAGGCGCTTCCA 2198  
QY 901 AGTTGATGCCCTTCTCTCT---TTCTCTCAATCAATGTCAGGAGTCAAAAGGCTGT-- 955  
Db 2199 AGATGACACCCCTTCT 2258  
QY 956 --AGCACAGAGTGGAGTTGATTAT 979  
Db 2259 ACAGCACAGATGGAATTTGATTAT 2284  
RESULT 9  
BM802135  
LOCUS  
DEFINITION  
AGENCOURT\_6461654 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5559503  
5', mRNA sequence.  
BM802135  
BM802135.1 GI:19118958  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 957)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1284 row: c column: 24  
High quality sequence stop: 621.  
FEATURES  
Location/Qualifiers  
1. 957  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

```
/clone="IMAGE:S559503"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/notes="Organ: Small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

BASE COUNT      243 a      240 c      246 g      224 t      4 others
ORIGIN

Query Match      66.7%; Score 764.6; DB 12; Length 957;
Best Local Similarity 97.9%; Pred. No. 3.3e-28;
Matches 796; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

QY 219 AAAATCATCATCCCGGCTGAGAGGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATTG 278
Db 10 AAAATCATCATCCCGGCTGAGAGGCGCTTCAAGGTTTTCGAGCAGCTTGGGACATTG 69

QY 279 TCAGCAATGCTGAAAGCGAAAGAGTATGAGATGAAACGAATGGCAGAGAATGAGCTGA 338
Db 70 TCAGCAATGCTGAAAGCGAAAGAGTACGAGATGAAACGAATGGCAGAGAATGAGCTGA 129

QY 339 GCCGGTCAGTAAATGAGTTTCTGTCACAGCTGCAAGATGACCTCAAGGAGCAATGAATA 398
Db 130 GCCGGTCAGTAAATGAGTTTCTGTCACAGCTGCAAGATGACCTCAAGGAGCAATGAATA 189

QY 399 CTATGATGCTAGCCGATGCCAGGAAAGCATAGGAGTTTGAATGGACCGGACCTA 458
Db 190 CTATGATGCTAGCCGATGCCAGGAAAGCATAGGAGTTTGAATGGACCGGACCTA 249

QY 459 AGAGTGCACATAGTGTGCTGAGTGTAAATAGGTGTCATCTGCTGAGGAAGGAGACTTTT 518
Db 250 AGAGTGCACATAGTGTGCTGAGTGTAAATAGGTGTCATCTGCTGAGGAAGGAGACTTTT 309

QY 519 GGCAGAGTCAAGATGTTGGGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGG 578
Db 310 GGCAGAGTCAAGATGTTGGGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGG 369

QY 579 TGTATGACATACAGAGTGGCTGGATCCAGCGGTGTAGTATCTCCCGAGATACCCACA 638
Db 370 TGTATGACATACAGAGTGGCTGGATCCAGCGGTGTAGTATCTCCCGAGATACCCACA 429

QY 639 GAGTCCCTATCATCTCATTTGGTTCTCGGATTCAGGCACACAGAGGCGGCAGAGAG 698
Db 430 GAGTCCCTATCATCTCATTTGGTTCTCGGATTCAGGCACACAGAGGCGGCAGAGAG 489

QY 699 CCACCCAGATGCCCTCCTGCTGATCTTCAGGATTTCTTGAATCGGATCTTTCAAGTAC 758
Db 490 CCACCCAGATGCCCTCCTGCTGATCTTCAGGATTTCTTGAATCGGATCTTTCAAGTAC 549

QY 759 CCCAGGGCAGATGCCCAATGGGAATCTTTTTCAGCTCTCAGCTGCCCTGGAGCGG 818
Db 550 CCCAGGGCAGATGCCCAATGGGAATCTTTTTCAGCTCTCAGCTGCCCTGGAGCGG 609

QY 819 CTGACGCTCTAAGCCCAACAGACAGTACCCAGGGAGAGCCAACTTAAGCGGGGA 878
Db 610 CTGACGCTCTAAGCCCAACAGACAGTACCCAGGGAGAGCCAACTTAAGCGGGGA 669

QY 879 AGAAAGTGAAGGAGGCGCTTCCAAAGTTGATGCCCTTCTCTTTTCCTCAAAATCAATGTCAG 938
Db 670 AGAAAGTGAAGGAGGCGCTTCCAAAGTTGATGCCCTTCTCTTTTCCTCAAAATCAATGTCAG 729

QY 939 GGAGTCAAAAGGCGCTGAGCACA - GGATGAGTTTGAATTAAT - CCTCTCTCCGCCCAACA 995
Db 730 GGAGTCAAAAGGCGCTGAGCACAAGGATGGAGTTTGAATTAATCCTCTCTCCGCCCAACA 789

QY 996 CCTAGGAAGTGAATCTTTTCTTTTATTTTTT 1028
Db 790 CCTAGGAAGTGAATCTTTTCTTTTATTTTTT 822
```

```
RESULT 10
BG763500
LOCUS
DEFINITION
602735591F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4860694 5',
mRNA sequence.
BG763500
ACCESSION
BG763500.1 GI:14074153
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 878)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rsb@pilot.nih.gov
Tissue Procurement: ATCC/DCTD/PTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1718 row: f column: 23
High quality sequence stop: 811.
Location/Qualifiers
1..878
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4860694"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 227 a 211 c 238 g 201 t 1 others
ORIGIN
```

```
Query Match      66.2%; Score 759.4; DB 12; Length 878;
Best Local Similarity 95.1%; Pred. No. 6.1e-28;
Matches 837; Conservative 0; Mismatches 37; Indels 6; Gaps 5;

QY 263 GCAGCTTGGGACATTTGTCAGCAATGCTGAAAGCGAAAGAGTATGAGATGAAACGAATG 322
Db 2 GCAGCTTGGGACATTTGTCAGCAATGCTGAAAGCGAAAGAGTATGAGATGAAACGAATG 61

QY 323 GCAGAGATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGACCTC 382
Db 62 GCAGAGATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGACCTC 121

QY 383 AAGGAGCAATGAATACTATGATGTAGCCCATGCCAAGGAAGCATAGGAGTTTGA 442
Db 122 AAGGAGCAATGAATACTATGATGTAGCCCATGCCAAGGAAGCATAGGAGTTTGA 181

QY 443 ATGGACCGGGAACCTTAAGAGTCCAGATACTGTGCTGAGTGTAAATAGGCTGATCCTGCT 502
Db 182 ATGGACCGGGAACCTTAAGAGTCCAGATACTGTGCTGAGTGTAAATAGGCTGATCCTGCT 241

QY 503 GAGGAGGAGACTTTTGGGCGAGTCAAGCATGTTGGGCTTCAAGATCACCTACTTTGCA 562
Db 242 GAGGAGGAGACTTTTGGGCGAGTCAAGCATGTTGGGCTTCAAGATCACCTACTTTGCA 301

QY 563 CTGATGGATGGAAGGTGTATGACATCAAGAGTGGGCTGGATGCCAGCGTGTAGGTATC 622
```

```

RESULT 11
BU538351
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU538351
AGENCOUT 10181238 NIH_MGC 107 Homo sapiens cDNA clone
IMAGE:6568891 5', mRNA_sequence.
BU538351
BU538351.1 GI:22948792
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 955)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM2755 row: e column: 19
High quality sequence stop: 583.
Location/Qualifiers
1..955
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6568891"
FEATURES
source

```

directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH MGC Library."

BASE COUNT	219 a	244 c	239 g	202 t
ORIGIN				
Query Match	60.4%	Score 692.8;	DB 13;	Length 904;
Best Local Similarity	95.5%;	Pred. No. 7.6e-25;		
Matches	745;	Conservative	0;	Mismatches 32; Indels 3; Gaps 3;
QY	254	GTCTTGGAGCAGCTTGGGACATGTGTCAGCAATCTCAGAAAGCGAAGGAGTATGAGATG	313	
Db	1	GTCTTGGAGCAGCTTGGGACATGTGTCAGCAATCTCAGAAAGCGAAGGAGTATGAGATG	60	
QY	314	AAACGAATGCGAGAGATGAGCTGAGCGGTGACGTAATGAGTTCTGTCCAAGCTCAA	373	
Db	61	AAACGAATGCGAGAGATGAGCTGAGCGGTGACGTAATGAGTTCTGTCCAAGCTCAA	120	
QY	374	GATGACCTCAAGGAGGCAATGAATATCTATGATGTGTAGCCGATGCCAAGGAAGCATAGG	433	
Db	121	GATGACCTCAAGGAGGCAATGAATATCTATGATGTGTAGCCGATGCCAAGGAAGCATAGG	180	
QY	434	AGTTTGAATGACCGGGAACTTAAGAGTGCAGATCTGTGCTGATGTATAGGCTG	493	
Db	181	AGTTTGAATGACCGGGAACTTAAGAGTGCAGATCTGTGCTGATGTATAGGCTG	240	
QY	494	CATCTGCTCAGGAAGAGAGACTTTTGGGCGAGTCAAGCATGTGGGCTCAAGATCACC	553	
Db	241	CATCTGCTCAGGAAGAGAGACTTTTGGGCGAGTCAAGCATGTGGGCTCAAGATCACC	300	
QY	554	TACTTTGCACTGATGGTGAAGAGGTGTATGATCATCACAGATGGGCTGGATGCCAGCT	613	
Db	301	TACTTTGCACTGATGGTGAAGAGGTGTATGATCATCACAGATGGGCTGGATGCCAGCT	360	
QY	614	GTAGGTATCTCCGAGATACCCACAGAGTCCCTATCATCTCATTTGGTTCTCGGATT	673	
Db	361	GTAGGTATCTCCGAGATACCCACAGAGTCCCTATCATCTCATTTGGTTCTCGGATT	420	
QY	674	CCAGGCACAGAGGGCGGAGAGAGCCACCCAGATGCCCTCTGCTGATCTTCAGGAT	733	
Db	421	CCAGGCACAGAGGGCGGAGAGAGCCACCCAGATGCCCTCTGCTGATCTTCAGGAT	480	
QY	734	TTCTTGAAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGGGAATCTTTTGA	793	
Db	481	TTCTTGAAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGGGAATCTTTTGA	540	
QY	794	GCTCCTCAGCCTGCCCTGGAGCGGTGTCAGCCTCTAAGCCCAACAGCAGTACCCAAAG	853	
Db	541	GCTCCTCAGCCTGCCCTGGAGCGGTGTCAGCCTCTAAGCCCAACAGCAGTACCCAAAG	600	
QY	854	GGAGAAGCCAAACCTAAGCGGCGAAGAAAGTGAAGGAGG-CCCTTCCAAGCTTGAAGCC	912	
Db	601	GGAGAAGCCAAACCTAAGCGGCGAAGAAAGTGAAGGAGGCTTCCAAGCTTGAAGCC	660	
QY	913	CTTCTCTTCTCAATCATGTGAGGAGTCAAAAGGCTGTAGCA-CAGGATGAGATT	971	
Db	661	CTTCTCTTCTCAATCATGTGAGGAGTCAAAAGGCTGTAGCA-CAGGATGAGATT	720	
QY	972	T-GAATTTATCCCTCTCCCTCCCTCCCTTACCTAAGGAAGTCACTCCCTCTTCCCTTTTGA	1030	
Db	721	TGAATTTATCCCTCTCCCTCCCTCCCTTACCTAAGGAAGTCACTCCCTCTTCCCTTTTGA	780	

RESULT 14  
BM452174  
LOCUS  
DEFINITION  
AGENCOURT\_6386043 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5526530  
5', mRNA sequence.  
ACCESSION  
BM452174  
VERSION  
BM452174.1 GI:18501214  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 995)  
NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: rgs@nih.gov  
Tissue Procurement: ATCC/DCTD/DMF  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL2200 row: f column: 03  
High quality sequence stop: 638.  
Location/Qualifiers  
1..995  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5526530"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_72"  
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

BASE COUNT 224 a 273 c 249 g 248 t 1 others  
ORIGIN  
Query Match 58.6%; Score 672.2; DB 12; Length 995;  
Best Local Similarity 99.6%; Pred. No. 6.5e-24;  
Matches 674; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 471 ACTGTGCTGAGTGTATAGCTGCATCTCTGCTGAGGAAGAGAGCTTTTGGGACAGTCAA 530  
Db 1 ACTGTGCTGAGTGTATAGCTGCATCTCTGCTGAGGAAGAGAGCTTTTGGGACAGTCAA 60  
QY 531 GCATGTTGGGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTATGACATCA 590  
Db 61 GCATGTTGGGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTATGACATCA 120  
QY 591 CAGAGTGGGCTGATGCCAGCTGAGTATCTCCGAGATACCCACAGATCCCTTATC 650  
Db 121 CAGAGTGGGCTGATGCCAGCTGAGTATCTCCGAGATACCCACAGATCCCTTATC 180  
QY 651 ACATCTCATTTGGTTCTCGGATTCAGGACACAGAGGGCGGAGAGAGCCAGATG 710  
Db 181 ACATCTCATTTGGTTCTCGGATTCAGGACACAGAGGGCGGAGAGAGCCAGATG 240  
QY 711 CCCTCTCTGCTGATCTTCAAGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGAG 770  
Db 241 CCCTCTCTGCTGATCTTCAAGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGAG 300  
QY 771 TGCCCAATGGGAATCTTTTGGAGCTCTCAGCTGCGCTCGAGCGCTGACGCTCTA 830  
Db 301 TGCCCAATGGGAATCTTTTGGAGCTCTCAGCTGCGCTCGAGCGCTGACGCTCTA 360  
QY 831 AGCCCAACAGCAGTACCCCAAGGGAGAGCCAAAGCTTAAGCGCGGAGAAAGTGAAG 890  
Db 361 AGCCCAACAGCAGTACCCCAAGGGAGAGCCAAAGCTTAAGCGCGGAGAAAGTGAAG 420  
QY 891 GGCCCTTCCAAGTTGATGCCCTTCTCTTTCTCAATCAATGTGAGGAGTCAAAAGG 950  
Db 421 GGCCCTTCCAAGTTGATGCCCTTCTCTTTCTCAATCAATGTGAGGAGTCAAAAGG 480  
QY 951 GCTGTAGCAGAGATGGAGTTGATTTATCCCTCTCCCTCCCTCCCTAGGAATCAATC 1010  
Db 481 GCTGTAGCAGAGATGGAGTTGATTTATCCCTCTCCCTCCCTCCCTAGGAATCAATC 540





QY	541	CCTCAGATCACCTACTTTCACCTGATGGATGGAAAGTGTATCACATCACAGAGTGGC	600
Db	110652		
QY	601	TGGATGCCACGCTGTAGGTATCTCCCCAGATACCCACAGAGTCCCTTATCACATCTCATT	660
Db	110712		
QY	661	TGGTTCGCGATTCCAGGCAACAGAGGCGGACAGAGCCACCCAGATGCCCTCTCTGC	720
Db	110758		
QY	721	TGATCTTCAGGATTCTTGAGTCGGATCTTTCAAGTACCCCGGCGAGATGCCCAATGG	780
Db	110818		
QY	781	GAACTTCTTTGCAGCTCCTCAGCCTGCCCTGGAGCGGTGCAGCCTCTAAGCCCAACAG	840
Db	110878		
QY	841	CACAGTACCCAGGGAGAGCCAAAACCTAAGCGGCGAAGAAAGTGAAGAGGCCCTTCCA	900
Db	110938		
QY	901	AGTTGATGCCCTT---CTCTTCTCAATCAATGTCAAGGAGTCAAAAGGCTGT--	955
Db	110998		
QY	956	--AGCACAGGATGGAGTTTGATTAT	979
Db	111058		
		ACAGCACGGGATAGAAATTTGATTAT	111083

Search completed: December 1, 2003, 08:50:25  
Job time : 4382 secs

Lorensuhera, L., Loulaged, H., Lozado, R.J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwakeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pascernak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puzo, M., Quitoz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,  
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
Weinstock, G. and Gibbs, R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 247309)  
Worley, K.C.  
Direct Submission  
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 247309)

Rat Genome Sequencing Consortium.  
Direct Submission  
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 10, 2003 this sequence version replaced gi:25012326.

The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GPIX  
Center clone name: CH230-57L7  
----- Summary Statistics  
Assembly program: Atlas 3.0;  
Consensus quality: 23539 bases at least Q40  
Consensus quality: 23537 bases at least Q30  
Consensus quality: 236797 bases at least Q20  
Estimated insert size: 246711; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 246159: contig of 246159 bp in length  
\* 246160 246259: gap of unknown length  
\* 246260 247309: contig of 1050 bp in length.  
FEATURES  
Source  
1..247309  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-57L7"  
1..1645  
/note="wgs end\_extension"  
clone\_end:Sp6  
6457..7267  
/note="clone boundary"  
clone\_end:Sp6  
site:EcoRI  
end\_sequence:BH299097"  
239013..239869  
/note="clone boundary"  
clone\_end:T7  
site:EcoRI  
end\_sequence:BH299095"  
243192..246159  
/note="wgs end\_extension"  
clone\_end:T7"  
BASE COUNT 69270 a 46265 c 47122 g 74568 t 10084 others  
ORIGIN

Query Match 57.6%; Score 661.2; DB 2; Length 247309;  
Best Local Similarity 81.8%; Pred. No. 3.3e-202;  
Matches 807; Conservative 0; Mismatches 158; Indels 21; Gaps 3;  
QY 1 GCCTTGGGTCAGCAGAGATATTTAATAGGCGGGGAATGCACCTGTAGTGTGGGGCTTA 60  
DB 110112 GCTCTGAGTCATGCGAGAGACTAAGAAACAGAGCAATGCCCTGTAGTGGGAGATA 110171  
QY 61 CTGCCAGCTGAGAGGAGTGGCTCGACTCTTGACCATTGGTGGGTTCTCTGAGATGA 120  
DB 110172 CTGCAGCGCTGAGAGGAGTAGCCCGACTCTTGATTATGGTGGGTTCTCTGAGATGA 110231  
QY 121 GCTAAACCTTTCCATGTACTGGGGTTGAGCCACAGCATCAGATGTTGAACTGAAGA 180  
DB 110232 ACTAAACCTTTTCAATGTCTAGGGTTGAGCTACAGCTCCGATATTGAACTGAAGA 110291  
QY 181 GGCCTATAGCAGCTGGCAGTGATGGTTTCATCTGCACAAAAATCATCCCCGGGCTGA 240  
DB 110292 GGCCTATAGGCAACTAGTGGTAATGATCCATCTGATAAAATCACATCCCAAGGCTGA 110351  
QY 241 GGAGCCCTTCAAGGTTTTCGAGCAGCTGGGACATTTGTAGCAATGTGAAAAAGGAAA 300  
DB 110352 GGAGCCCTTCAAGGTTTTCGCGGAGCTTGGGACATTTGTAGCAATGTGAAAAAGGAAA 110411  
QY 301 GGAGTATCAGATGAACGAATGGCAGAGAAATGAGCTAGCCGGTCAAGTAATAGTTTCT 360  
DB 110412 GGAATATGAGATAAAATGAATGGCAGAGAGTAGCTACGCCGGTTAGTGAATAGTTTCT 110471  
QY 361 GTCCAAGCTGCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTAGCCGATGCCA 420  
DB 110472 GTCCAGAGCTACAGATGACCTCAAGAGGCAATGAATACTATGATGTAGCCGATGCCA 110531  
QY 421 AGGAAAGCATAGGAGGTTTGAATGGACCGGGAACCTAAGAGTGCAGATCTGTGCTGA 480  
DB 110532 AGGAAAGCATAGGAGGTTTGAATGGAGGAGTGCAGATCTGTGCTGA 110591  
QY 481 GTGTAATAGGCTGATCTCTGAGGAGGAGATCTTTGGCAGAGTCAAGCATGTGGG 540  
DB 110592 CTGTAACAGGGCTTCATCTGCTGAGGAAAGAGTCTTCTGGGTAGAGTCAAGCATGTGGG 110651

QY 901 ACGTTGATGCCCTTCTCTT--TCTCAAAATCAATGTCAAGGAGTCAAAAGGCTGT--- 955

Db 142449 ACGATGAAACCACTTCTCTCTCTCTTCTCAGTCAATATCAGGAGTCAAAAGGCTGTGTA 142390

QY 956 -AGCACAGGATGGAGTTTGATTAT 979

Db 142389 CAGCACAGGATGGAGTTTGATTAT 142365

RESULT 14

AF141342

LOCUS AF141342 754 bp mRNA linear PRI 14-JAN-2001

DEFINITION Homo sapiens LYST-interacting protein Lip6 mRNA, complete cds.

ACCESSION AF141342

VERSION AF141342.1 GI:12239359

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 754)

AUTHORS Tchernev,V.T., Mansfield,T.A., Giot,L., Kumar,A.M., Mandabalan,K., Li,Y., Mishra,V.S., Detter,J.C., Rothberg,J.M., Wallace,M.R., Southwick,F.S. and Kingsmore,S.F.

TITLE Interactions of the Chediak-Higashi lysosomal-trafficking regulator protein with SNARE complex and signal transduction proteins

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 754)

AUTHORS Tchernev,V.T., Mansfield,T.A., Giot,L., Kumar,A.M., Mandabalan,K., Li,Y., Mishra,V.S., Detter,J.C., Rothberg,J.M., Wallace,M.R., Southwick,F.S. and Kingsmore,S.F.

TITLE Direct Submission

JOURNAL Submitted (08-APR-1999) Full-Length Cloning, CuraGen Corporation, 12085 Research Drive, BDI, Alachua, FL 32615, USA

FEATURES

source 1..754

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

11..586

/codon\_start=1

/product="LYST-interacting protein Lip6"

/protein\_id="AAG49445.1"

/db\_xref="GI:12239360"

/translations="MKRMAENELSRVNEFLQLDLEAMNTMCSRQCKHRRFE  
MOREPKSARYCABCNLHPAEEGDFAWESMLGKITYFLMDGKYVDITEWAGCGRV  
GISPDHRYVPHLSFGSRIPGTRGRORAFDAPADLQDLFSRIFVPPQMPMGTSLS  
QLLSLELPELPLSPSLPAQVPREPNLSGGRK"

BASE COUNT 193 a 185 c 200 g 176 t

ORIGIN

Query Match 64.8%; Score 743; DB 9; Length 754;

Best Local Similarity 99.9%; Pred. No. 5.6e-229;

Matches 754; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 301 GGAGTATGAGTGAACGAATGCCAGAGATGAGCTGAGCGGTCACTAAATCAGTTTCT 360

Db 1 GGAGTATGAGTGAACGAACGAATGCCAGAGATGAGCTGAGCGGTCACTAAATCAGTTTCT 60

QY 361 GTCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATCTATGATGTAGCCGATGCCA 420

Db 61 GTCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATCTATGATGTAGCCGATGCCA 120

QY 421 AGGAAGCATAGAGGTTTGAATGACCGGGAACCTTAAGAGTCCAGATCTGTGCTGA 480

Db 121 AGGAAGCATAGAGGTTTGAATGACCGGGAACCTTAAGAGTCCAGATCTGTGCTGA 180

QY 481 GTGTAATAGGCTGCATCTCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTGG 540

Db 181 GTGTAATAGGCTGCATCTCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTGG 240

QY 541 CCTCAGATCCTACTTTGCACTGATGGATGGAAGGTGTATGATCACAGAGTGGCC 600

Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Wainstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Unpublished  
2 (bases 1 to 263371)  
Worley,K.C.

TITLE  
JOURNAL

Direct Submission  
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 263371)  
Worley,K.C.

TITLE  
JOURNAL

Direct Submission  
Submitted (10-OCT-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Oct 10, 2002 this sequence version replaced gi:21953452.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: TUDU

Center clone name: CH230-2H10

----- Summary Statistics

Assembly program: Phrap; version 0.930329

Consensus quality: 243674 bases at least Q40

Consensus quality: 246830 bases at least Q30

Consensus quality: 248548 bases at least Q20

Estimated insert size: 247731; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 258056: contig of 258056 bp in length

\* 258057 258156: gap of unknown length

\* 258157 261384: contig of 3228 bp in length

\* 261385 261484: gap of unknown length

\* 261485 263371: contig of 1887 bp in length.

FEATURES

source

1..263371  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-2H10"

misc\_feature

1..1125

/note="wgs\_contig"

misc\_feature

2121..4146

/note="wgs\_contig"

misc\_feature

5480..6935

/note="wgs\_contig"

misc\_feature 256941..258056  
/note="wgs\_contig"  
misc\_feature 258157..259735  
/note="wgs\_contig"  
misc\_feature 259889..261384  
/note="wgs\_contig"  
BASE COUNT 75390 a 46168 c 46816 g 81072 t 13925 others  
ORIGIN

Query Match 66.8%; Score 765.8; DB 2; Length 263371;  
Best Local Similarity 87.5%; Pred. No. 4e-236;  
Matches 862; Conservative 0; Mismatches 117; Indels 6; Gaps 2;

QY 1 GCCTTCGGGTCAACAGCAATATTAAATAGGAGAGGGAATGACCTGTAGCTAGTGGGCCCTA 60  
DB 143349 GCCTTCGGGTCAACAGAGGACTCAGACACAGGCACTGCCCATGTAAATACGCGCGATA 143290  
QY 61 CTGCCAGCCTTGAAGAGGAAGTGGCTCGACTCTTTCGACCATGGCTGGGGTTCCTCAGGATGA 120  
DB 143289 CTGCCAGCCTTGAAGAGGAAGTGGCTCGACTCTTTCGACCATGGCTGGGGTTCCTCAGGATGA 143230  
QY 121 GCTAAACCCCTTCCATGTACTCTGGGGTTGAGGCCACAGCATCAGATGTTTCAACTGAAGAA 180  
DB 143229 ACTAAACCCCTTTCATGTCTGGGGTTGAAGCTACAGATCCGATATTGAACTAAAGAA 143170  
QY 181 GGCCTATAGACAGCTGGCAGTAGTGTTCATCTCTGACAAAAATCATCATCCCCGGGCTGA 240  
DB 143169 AGCCTATAGCAGCTAGCGGTGATGCTCCATCTGATATAAAATCAACATCCCCGGGCTGA 143110  
QY 241 GGAGGCGCTTCAAGGTTTTCGGAGCAGCTTGGGACATTTGAGCAATGCTGAAAAAGGAAA 300  
DB 143109 GGAGGCGCTTCAAGGTTTTCGGGCGAGCTTGGGACATTTGAGCAATCCAGAGAGCGGAA 143050  
QY 301 GGAGTATGAGATGAACAGAAATGGCAGAAATGAGCTGAGCGGTGAGTAAATCAGTTTCT 360  
DB 143049 GGAATATGAGATGAACAGGATGGCAGAAATGAGCTCAGCGCTCGTGAATGAGTTTCT 142990  
QY 361 GTCCAAGCTGCAAGATGACCTCAAGAGGCAATGAATATGATGTGAGCGGATGCCA 420  
DB 142989 TTCGAAGCTACAGATGATTTCAAGAGGCAATGAATGATGATGTCAGCGCGTGCCA 142930  
QY 421 AGGAAGCATAGAGGTTTGAATGAGCGGAACTTAAGAGTGCAGATATCTGTCTGA 480  
DB 142929 AGGGAAGCATAGAGGTTTGAATGAGCGGAACTTAAGAGTGCAGATATCTGTCTGA 142870  
QY 481 GTGTAATAGCTGTCATCTGCTGAGGAAGGACATTTTGGGAGAGTCAAGCATGTTGG 540  
DB 142869 GTGTAAACAGGTTGTCATCTGCTGAGGAAGGAGACTTCTGGGAGAGTCAAGCATGTTGG 142810  
QY 541 CCTCAAGATCACTACTTTGCACTGATGATGGAAGGTGATGACATCAGAGTGGGC 600  
DB 142809 CCTCAAGATCACTACTTTGCACTGATGATGGAAGGTGATGACATCAGAGTGGTC 142750  
QY 601 TGGATGCCAGCGTGTAGTATCTCCCAAGATACCCACAGAGTCCCTCATCATCTCAIT 660  
DB 142749 TGGATGCCAGCGTGTGGGCATCTCCCAAGATACCCACAGAGTCCCTCATCTCAIT 142690  
QY 661 TGGTTCTCGGATTCAGGCAACAGAGGCGGAGAGACCCAGATGCCCTCTCTGC 720  
DB 142689 TGGTTCTCGGTTACCAGGCAACAGGCGGCGGACAAACAGCACTCCAGAGTCCCTCTCTGC 142630  
QY 721 TGATCTTCAGGATTTCTTTGAGTCGGATCTTTCAAGTACCCCGAGGAGAGTCCCAATGG 780  
DB 142629 TGACCTGAGGATTTCTTTGAGTCGGATCTTTCAAGTACCTCCAGGCGGATGTCCAATGG 142570  
QY 781 GAACTTTCTTCAGCTCTCCTCAGCCTGCGGAGCGGCTGAGCGCTCTAGAGCCCAAG 840  
DB 142569 GAACTTTCTTCAGCTCTCCTCAGCCTGCGGAGCGGCTGAGCGCGGAGCGGCTTCCA 142510  
QY 841 CACAGTACCCAGGAGAGCAACCTTAAGCGGCGGAGAGAAAGTGAAGAGCGGCTTCCA 900  
DB 142509 TTCAGTATCCAGGAGAGCAACCTTAAGCGTGGAGAGAGTGAAGTGAAGTGGCGGCTTCCA 142450

```
/gene="Drip78"  
/note="Drip78; ER-membrane protein; zinc finger motif"  
/codon_start 1  
/product="dopamine receptor interacting protein"  
/protein_id="AAK56240.1"  
/db_xref="GI:14194055"  
/translation="MAQKHPGERGLCGVHSHSGSSLITSSGSDPEILSPSGLRDSKE  
TAPNGTRCLKEHSDPKCTQSPNPAHSDPSHPGPPREGGYDSESTCEBEGV  
DQLSRENEYQDGSFSLP;PSACNCGSPGVPGTCSERGDSSSFCHCTSP  
ALGEDSELEBYDEPLKFPSPFSRVSSGKPPSRQRHRLTKEDVRDGRDPKA  
FGHRLAEVQITPEGLGKMGVEGLQGLQGAGFWLIELLVGEYVETCGYLIYACR  
KLKSGDLDFRIWGVNWARLGGMARVMFQLSQSPFVAGLFIRLLVVGAFLIAL  
ALFLGLQLGRLVGLDRLGWRKAAMFLMDSPALHFLTLTKDSPWQOLYV  
IOWMLELPMVKORTQGTANVSGRYCOPREVEARLLTMAGVPEDELNPFVLGVE  
ATASDLELKAVQLAMVHPDKNHHPRAEEAFKLRAMWDIVSNPERRKEYEMKMA  
ENELSVNEFLSKODLKAANNWCMSCQKRRFEMDREPKFARCAECNRLHPA  
AEGEDWASSMGLKITITFALMDKVDITWAGQQRVGSPTDHPVPHVPSFGRV  
PGTSGRATPEPPADLDQFLSLRIFVQPPGPMNGNFFAAPHGPGLTTSRPNSSV  
PKGEAKPKRKKRRRPFQR"  
BASE COUNT 565 a 582 c 723 g 522 t  
ORIGIN
```

```
Query Match 68.2%; Score 781.8; DB 10; Length 2392;  
Best Local Similarity 88.5%; Pred. No. 1.7e-241;  
Matches 872; Conservative 0; Mismatches 107; Indels 6; Gaps 2;  
  
QY 1 GCCTTGGGTCAACGAGAAATTAATAGCGAGGGAATGCACCTGTAGCTAGTGGCGCTA 60  
Db |||||  
QY 1360 GCCTGGGTCAACGAGGACTCAGAGGAGGCACTGCCTGCTAGCTAGCGGGGATA 1419  
Db |||||  
QY 61 CTGCCAGCTGAAGAGGAGTGCTCGACTCTTGACCATGGCTGGGGTCTCTGAGATGA 120  
Db |||||  
QY 1420 CTGCCAGCTGAAGAGGAGTGCTCGACTCTTGACCATGGCGAGGGGTTCTTGAAGATGA 1479  
Db |||||  
QY 121 GCTAAACCTTTCCATGTACTGGGGTGTGAGGCGACAGCATCAGATGTTGAACTGAAGA 180  
Db |||||  
QY 1480 ACTAAACCTTTTCATGTCTGGGGTGTGAGCTACAGATCGATATTGAACTAAAGAA 1539  
Db |||||  
QY 181 GGCCTATAGACAGCTGGCAGTATGTTTCATCTGTGACAAATATCATCTCCCGGGCTGA 240  
Db |||||  
QY 1540 GGCCTATAGCAGCTAGCGGTGATGTGCTTCTGATAAAACCCACCATCCCGGGCTGA 1599  
Db |||||  
QY 241 CGAGGCTTCAAGGTTTTCGAGGAGCTTGGGACATTTGCAGCAATGTGCAAAAGCGAA 300  
Db |||||  
QY 1600 GGAGGCTTTAAAGTTTTCGGGAGCTTGGGACATCTGTCAGCAATCCAGAGGCGGAA 1659  
Db |||||  
QY 301 CGAGTATGAGATGAACCAATGGCAGAGAAATGAGCTGAGCGGTGAGTAAATGAGTTTCT 360  
Db |||||  
QY 1660 GGAATATGAGATGAACCGATGGCAGAGATGAGCTCAGCCCTCGGTGAATGAGTTTCT 1719  
Db |||||  
QY 361 GTCCAGCTGCAAGATGACCTCAAGGAGCAATGAATACTATGATGTGTAGCGGATGCCA 420  
Db |||||  
QY 1720 GTCCAGCTGCAAGATGACCTCAAGGAGCAATGAATAGATGATGTGAGCGGATGCCA 1779  
Db |||||  
QY 421 AGGAAGCATAGGAGTTTGAATGAGCGGGAACCTTAAGAGTGCAGATCTGTGCTGA 480  
Db |||||  
QY 1780 AGGAAGCATAGGAGTTTGAATGAGCGGGAACCTTAAGAGTGCAGATCTGTGCTGA 1839  
Db |||||  
QY 481 GTGTAATAGGCTGATCTGCTGAGGAGGAGACATTTTGGGAGAGTCAAGCATGTTGGG 540  
Db |||||  
QY 1840 GTGTAACAGGTTGATCTGCTGAGGAGGAGACATTTTGGGAGAGTCAAGCATGTTGGG 1899  
Db |||||  
QY 541 CCTCAAGATCACTACTTTGACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGGC 600  
Db |||||  
QY 1900 CCTCAAGATCACTACTTTGACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGGC 1959  
Db |||||  
QY 601 TGGATGCGAGCTGTAGTATCTCCCGAGATACCCACAGAGTCCCTTACATCTCAT 660  
Db |||||  
QY 1960 TGGATGCGAGCTGTAGTATCTCCCGAGATACCCACAGAGTCCCTTACATCTCAT 2019  
Db |||||  
QY 661 TGGTCTCGGATTCAGGACCAAGGGCGCAGAGCCACCCAGATGCCCTCTCTGC 720  
Db |||||  
QY 2020 TGGTCTCGGATTCAGGACCAAGGGCGCAGAGCCACCCAGATGCCCTCTCTGC 2079  
Db |||||
```

```
QY 721 TGATCTTTCAGGATTTCTTGATCGGATCTTTCAGTACCCCGAGGCGAGATGCCAATGG 780  
Db |||||  
QY 2080 TGACCTGCGAGATTTCTTGAGCGGATCTTTCAGTACCTCCAGGCGCGATGTCATGG 2139  
Db |||||  
QY 781 GAATCTTTCAGGATTTCTTGATCGGATCTTTCAGTACCTCCAGGCGCGATGTCATGG 840  
Db |||||  
QY 2140 GAATCTTTCAGGATTTCTTGATCGGATCTTTCAGTACCTCCAGGCGCGATGTCATGG 2199  
Db |||||  
QY 841 CACAGTACCCCAAGGAGAGCCAAACCTTAAGCGCGGAGAGAAAGTGAAGGCGGCTTCCA 900  
Db |||||  
QY 2200 TTCAGTACCCCAAGGAGAGCCAAACCTTAAGCGCGGAGAGAAAGTGAAGGCGGCTTCCA 2259  
Db |||||  
QY 901 ACCTTTCAGGATTTCTTGATCGGATCTTTCAGTACCTCCAGGCGCGATGTCATGG 955  
Db |||||  
QY 2260 AAGATGACACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2319  
Db |||||  
QY 956 -AGCACAGGATGAGTTGATTTAT 979  
Db |||||  
QY 2320 CAGCGCAGGATGAGTTGATTTAT 2344  
Db |||||
```

```
RESULT 13  
AC098454/c  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-2H10, WORKING DRAFT SEQUENCE, 3  
unordered pieces.  
AC098454  
AC098454.4 GI:23664557  
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.  
Rattus norvegicus (Norway rat)  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 263371)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
Albrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,  
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Geo, J., Garcia, A., Garner, T., Garza, N., Gall, R.,  
Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
Honsi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karissom, B., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratoch, J., Kuchta, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtharge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H.,  
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwodu, G.,  
Ogunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sison, I.,  
Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansy, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,  
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
```

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

FEATURES  
 Location/Qualifiers  
 1..2402  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CZECH II"  
 /db\_xref="taxon:10090"  
 /clone="MGC:19282 IMAGE:4016209"  
 /tissue\_type="Mammary tumor metastasized to lung."  
 MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR  
 enhancer."  
 /clone\_lib="NCI CGAP\_Lu30"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 148..2259  
 /codon\_start=1  
 /product="RIKEN cDNA 5730551F12 gene"  
 /protein\_id="AAH1146.1"  
 /db\_xref="GI:15029846"  
 /db\_xref="LocusID:74330"  
 /translation="MAQKHQPERRLCAHRSGGTSLSTSGSSVDPPEILSPFSGURDSAE  
 TAPNRCLEKHSQPKYTPPNPAHSDNPGHPGPPGPPGGYDPESTTGBEESGV  
 DOELSRNETGYDEGSGSFUSIPSCNQCQSPGVEGTYSDEGSSSLCHICTSP  
 ALGDELEEEYDDDEKLPFSDFSFSSVSSKKPLSRQKRRFPFKEDVROSGRREPKA  
 PGRRLARKRGLQDQKRGILGWVEELCOLQAGFWMLIELVLVGEYVETCGHLIYA  
 CRKLKSGDLDFRVMVWASRLGGMARMEQPLSQSFPCVVGILLRILKRVGALLL  
 ALALFLACQLQFWRVSVGLGNRLGRKWTALFWLGSPLALHCLTLKKDSRPWQVL  
 RLIIQWGLPEWPKRTKQKNQPAVSGRYQPEEVEVTRLTWAGVEDLNPFLVGL  
 VEATSDTELKAYRLQVAMPDKNHMPRAEEAFKILRAAWDITVSPERKKEVEMKR  
 MAENELSVNEFYLKLQDDLEAMTNMHCRCQGHRRFEMDRPKSAVCACENRL  
 HPABEGGFMAESSMLGLKITYFALMDGKAYDITWAGCQRVGI SPDTHRYPYHISFGS  
 RVPQTSQRORATPESPPVDLQDFLSRIFQVPPGPMNGNFFAAPHQPGGITTSRSPNS  
 SVPGAEKPKRRKVKVRPQR"  
 BASE COUNT 618 a 575 c 698 g 511 t  
 ORIGIN  
 Query Match 69.2%; Score 793.6; DB 10; Length 2402;  
 Best Local Similarity 89.2%; Pred. No. 2.5e-245;  
 Matches 880; Conservative 0; Mismatches 99; Indels 7; Gaps 2;  
 QY 1 GCCTTGGGTCAAGCAGAATATTATAGCGAGGGAATGCACCTGTAGCTAGTGGCGCTA 60  
 Db 1353 GCCCTGGGTCAAAACAGAGGACTAAGAAACAGGGCAATGCACCTGTAGCTAGCGGGCATA 1412  
 QY 61 CTGCACGCTTGAAGCAGAGTGGCTGCACCTTTGACCATGCTGGGTTCCTCAGGATGA 120  
 Db 1413 CTGCACGCTTGAAGCAGAGTGAACCCGACTCTTGACCATGGCTGGGGTTCTCTGAAGATGA 1472  
 QY 121 GCTAAACCTTTTCCATGTACTTGGGGGTGTAGGCCACAGCATCAGATGTTCAACTGAAGA 180  
 Db 1473 ACTAAACCTTTTTCATGCTGTGGGGTTGAAGCTACAGCATCCGACACTGAATAAAGAA 1532  
 QY 181 GGCTATAGACAGCTGCGAGTGATGGTTTCACTGACAAAAATCATCATCCCCGGGCTGA 240  
 Db 1533 GGCTATAGCAGCTAGCAGTAATGGTTCCTATCCCTGATAAAAATCACATCCCCGGGCTGA 1592  
 QY 241 GGAGGCCCTTCAAGTTTTTGGCAGCAGCTTGGGACATTGTCCAGCAATGCTGAAAGCGAAA 300  
 Db 1593 GGAGGCCCTTCAAAATTTTGGGGCAGCTTGGGACATTGTTCAGCAACCCAGAGGGCGAA 1652  
 QY 301 GGAGTATGAGATGAAACGAATGGCAGAGATGAGCTGAGCCGGTCACTAAATCAGTTTCT 360  
 Db 1653 GGAATATGAGATGAAACGGATGCGAGAGATGAGCTGAGCCGGTCACTGAATCAGTTTCT 1712  
 QY 361 GTCAGAGCTCAGATGACCTCAAGGAGGCAATGAAATACTATGATGTGAGCCGATGCCA 420  
 Db 1713 GTCGAAGCACTACAAGATGACCTCAAGGAGGCAATGAAACAGATGATGTGCAGCAGATGCCA 1772  
 QY 421 AGGAAAGCATAGAGAGTTTGAATGACCCGGGAACCTTAGAGTGCAGATCTGTGCTGTA 480  
 Db 1773 AGGAAAGCATAGAGAGTTTGAATGACCCGGGAACCAAGATGCGAGATCTGTGCTGTA 1832

BASE COUNT	618 a	575 c	698 g	511 t
ORIGIN				

Query Match	69.2%	Score 793.6	DB 10	Length 2402
Best Local Similarity	80.3%	Prod No	50345	

	Matches	880;	Conservative	0;	Mismatches	99;	Indels	7;	Gaps	2;
QY	1	GCCTTGGGTCAAGCAGAAATTAATAAGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA	60							
Db	1353	GCCTTGGGTCAAAACAGAGGACTAAGAAACAGGGCAATGCACCTGTAGCTAGCGGGCGATA	1412							
QY	61	CTGCCAGCCTGAAGCAGGAAGTGGCTCGACTCTTTGACCATGGCTGGGGTTCCTCAGGATGA	120							
Db	1413	CTGCCAGCCTGAAGCAGGAAGTGACCCGACTCTTTGACCATGGCTGGGGTTCCTCAGGATGA	1472							
QY	121	GCTAAACCCCTTTCATCTACTCTGGGGGTGAGGCCACAGCATCAGATGTTGAACCTGAAGAA	180							
Db	1473	ACTAAACCCCTTTCATGTGCTGGGGGTGAGGCTACAGCATCCGACACTGAATTAAGAA	1532							
QY	181	GGCCTATPAGACAGCTGGCAGTGATGGTTTCCTTGACAAAATCATCATCCCCGGGCTGA	240							
Db	1533	GGCCTATAGCAGCTAGCAGTAATGTTCCATCTGATAAAATTCACATCCCCGGGCTGA	1592							
QY	241	GGAGGCCTTCAAGGTTTTGCGCAGCCTTGGGACATTTGTGAGCAATGCTGAAAGCGAAA	300							
Db	1593	GGAGGCCTTCAAAATTTTGGCGGCAGCTTGGGACATTTGTGAGCAACCCAGAGAGGCGGAA	1652							
QY	301	GGAGTATGAGATGAACAGGAATGGCAGAGAAATGAGCTGAGCCCGTCACTAAATCAGTTTCT	360							
Db	1653	GGAAATGAGATGAAACGGATGGCAGAGAATGAGCTCAGCCGGTCACTGAATCAGTTTCT	1712							
QY	361	GTCCAAGCTGCAAGATGACCTCAAGCAGGCAATGAATACTATGATGTGAGCCGATGCCA	420							
Db	1713	GTCGAAACTACAAGATGAGCTCAAGGAGGCAATGAACACGATGATGTGCAGCAGATGCCA	1772							
QY	421	AGGAAAGCATAGAGAGTTTGAATGGACCGGGAACCTTAGAGTGGCCAGATACTGTGCTGA	480							
Db	1773	AGGAAAGCATAGAGAGTTTGAATGGACCGGGAACCCAAAGATGCCAGATACTGTGCTGA	1832							

Qy	481	GTGTAATAGGCTGCATCCTCTCGTAGGAAGAGAGACTTTTGGCGAGAGTCAAGCATGTGGG	540
Db	1833	GTGTAAACAGGCTGCATCCTGCTGAGGAAGAGAGACTTTTGGCGAGAGTCGAGCATGCTGG	1892
Qy	541	CCTCAAGATCACCTACTTTGCATGTGGATGGAAGGTGTATGACATCAAGAGTGGC	600
Db	1893	CCTCAAGATCACCTACTTTGCGTGTGGATGGAAAGGTGTATGACATTACAGAGTGGC	1952
Qy	601	TGGATGCCAGCGTGTAGGTATCTCCCAGATACCCACAGATGCCCTATCACATCTCAT	660
Db	1953	TGGATGCCAGCGTGTGGGTATCTCCCAGATATCTACAGAGTTCCITTACCACATCTCAT	2012
Qy	661	TGTTCTCGGATTCACAGCACAGAGCGGCGAGAGAGCCACCCACAGATGCCCTCTCTGC	720
Db	2013	TGTTCTCGGATACCCGGCACAGTGGCGGAGAGGGCCACTCCAGAGTCCCTCTCTGT	2072
Qy	721	TGATCTTCAGGATTTCTTGAGTCGGATCTTTTCAAGTACCCCCAGGCAGATGCCCAATGG	780
Db	2073	TGACCTGCGAGATTTCTTGAGCCGGATCTTTCAAGTACCTCGGGCCGAGTGTCCAATGG	2132
Qy	781	GAATCTTTTGAGCTCCTCAGCCTGCCCTGGAGCGCTGCAGCCTCTAAGCCCAACAG	840
Db	2133	GAATCTTTTGGCGCACTTCACTCGGCCCTGGAGCACTTTGACCTCTAGGCCCAACAG	2192
Qy	841	CACAGTACCCAAAGGAGAGCCAAACTAAGCGGGAAGAAAGTACAGAGGCCCTTCCA	900
Db	2193	TTCAAGTACCCAAAGGAGAGCCAAACTAAGCGGGAAGAAAGTACAGAGGCCCTTCCA	2252
Qy	901	ACGTTGATGCCCTTCTCT- - -TTCCTCAAAATCAATGTCAGGAGTCAAAAGGGCTGT- -	955
Db	2253	ACGATCACACCCCTTCTCTCTCTCTCAAAATCAATGTCAGGAGTCAAAAGGGCTGT	2312
Qy	956	--AGCACAGATGGAGTTTGATTTAT	979
Db	2313	ACAGCACAGATGGAAATTTGATTTAT	2338

RESULT 12  
AF351793

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS TITLE	JOURNAL
-------------------------------	------------------------------	-------------------------------	---------

[illegible]

gene CDS



MEDLINE 21436421  
PUBMED 11551646  
REFERENCE 2 (bases 1 to 2692)  
AUTHORS Neill,J.D. and Ridpath,J.F.  
TITLE Direct Submission  
JOURNAL Submitted (27-SEP-2000) VPDLRU, National Animal Disease Center,  
2300 Dayton Ave., Ames, IA 50010, USA  
REFERENCE 3 (bases 1 to 2692)  
AUTHORS Neill,J.D. and Ridpath,J.F.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-2001) VPDLRU, National Animal Disease Center,  
2300 Dayton Ave., Ames, IA 50010, USA  
REMARK Sequence update by submitter  
COMMENT On Oct 2, 2001 this sequence version replaced gi:12584131.  
FEATURES  
source  
1. 2692  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/notes="bDnaJ1"  
/codon\_start=3  
/product="DnaJ1 protein"  
/protein\_id="AAG59810.2"  
/db\_xref="GI:15843561"  
/translation="SAGSAPNGTRCTLHSHSPKPYTOPNPAPHWSDFSHPPRGPPPL  
AEDPDQSEASSSESGVDQELRENETGYDDNSFLSIPSCNQGPIDGEPYS  
EGRDSNSNFCRCHCTPALGEDEKLEJYDEELLSPFVNFHVCPEKXPAPRORPR  
VPAKETREGRDRPSRGRRLGRKSQADKRRGLMGAEELCOLGOAGFWLLEL  
LVLVGENVETCGHLIYACROLKSGDLDLSRWVGVWAGRLRCHQWFLSOGFCY  
AGLFTPLRLVGLALLLALLLGLCQLQHRFLARIEWTQGWDRDKATWIFSLASPT  
WQCLILLRESRQQLRVIQVWQELPVMKQRTNRQANPAGVGRYQCPREEVARL  
LTMWAGPEDELNFVGLVGEATASDELKAYQLAVMHPDKVHHPRAEEAFKVLRA  
ANDIVSNPRRKEYEYMKRVAENELSRVNEFLSKLOBAMNTVMCSRQCKGRSFFENDR  
EPKARYACENLHPAEEDRWAESMLGLKITFYALMDGKYVDITWAGQCRVCIS  
PDTHRVYHIISFSRMPGTSGRORATPDAPPADLQDFLSRIFQVPPFGMSNGNFFAAR  
QPGPGATAASKPNSTVPEKAPKRRKRRRPFQR"  
BASE COUNT 670 a 635 c 756 g 631 t  
ORIGIN  
Query Match 71.8%; Score 824; DB 4; Length 2692;  
Best Local Similarity 92.5%; Pred. No. 3.5e-255;  
Matches 882; Conservative 0; Mismatches 60; Indels 12; Gaps 1;  
QY 2 CCTTGGGTCAAGCAGAAATTAATAGCAGGGAATGCACCTAGCTAGCTAGTGGCGCTAC 61  
DB 1089 CTTTGGGTCAACAGAGGACCAATAGCAGGCGAATGCACCTAGCTAGCTAGTGGTCTTAC 1148  
QY 62 TGCACGCTGAAGAGGAAGTGGCTCGACTTTGACCATGCTGGGTTCTCAGGATGAG 121  
DB 1149 TGCACGCTGAAGAGGAAGTGGCTCGACTTTGACCATGCTGGGTTCTCAGGATGAG 1208  
QY 122 CTAAACCCCTTCCATGCTAGTGGGTTGAGGCGACAGCATCAGATTTGAACTGAAGAAG 181  
DB 1209 CTAAACCCCTTTCACGTGTGGGGTTGAGGCGACAGCATCAGATTTGAACTGAAGAAG 1268  
QY 182 GCCTATAGACAGCTGCGAGTGGTTCATCTGCACAAAATCATCATCCCGGCTGAG 241  
DB 1269 GCCTATAGCAGCTGCGAGTGGTTCATCTGCACAAAATCATCATCTCTGTCGGAG 1328  
QY 242 GAGGCTTCAAGTGTTCGAGCAGCTTGGGACATTTGTCAGCAATCTCTGAAAGGGAAG 301  
DB 1329 GAAGCTTCAAGTGTTCGCGGAGCTTGGGACATTTGTCAGCAATCTCTGAAAGCGGAAG 1388  
QY 302 GAGTATGATGAAGCAAGATGCGAGATGAGCTGAGCGGTGAGTAAATGAGTTCTG 361  
DB 1389 GAATATGATGAAGCAAGATGCGAGATGAGCTGAGCGGTGAGTAAATGAGTTCTG 1448  
QY 362 TCCAAGCTGCAAGATGACCTCAAGGAGCAATGAATACTATGATGTGAGCCGATGCCAA 421  
DB 1449 TCCAAGCTGCAA-----GAAGCAATGAATACCATGATGTGAGCGGATGCCAG 1496  
QY 422 GGAAGACATAGAGGTTTGAATGACCGGGAACCTAAGAGTGCCAGATCTGTGCTGAG 481

Db 1497 GGAAGCATAGGAGCTTTGAAATGGACCGGAACCTAAGAGTGCCAGATATCTGTGCTGAG 1556  
QY 482 TGTAAATAGGCTGATCTGCTGAGGAGGAGAGCTTTTGGGCAGAGTCAAGCATGTTGGGC 541  
Db 1557 TGTAAATAGGCTGATCTGCTGAGGAGGAGTGTCTTTGGGCAGAGTCAAGCATGTTGGGC 1616  
QY 542 CTCAGATACACCTACTTTTGACATGATGGATGGAAGGTGTATGACATCACAGAGTGGCT 601  
Db 1617 CTCAGATACACCTACTTTTGCGCTGATGGATGGAAGGTGTATGATATCACAGAGTGGCT 1676  
QY 602 GATGTCAGCGTGTAGTATCTCCCGAGATACCCACAGAGTCCCTTATCATCTCATTT 661  
Db 1677 GGATGTCAGCGTGTGGGAATCTCCCGAGATACCCACAGAGTCCCTTATCATCTCATTT 1736  
QY 662 GGTCTTCGAGTTCAGGACCAAGGCGGCGCAGAGAGCCACCCAGAGTCCCTCTCTGCT 721  
Db 1737 GGTTCAGGATGTCAGGACCAAGTGGGCGGCGCAGAGAGTCTACTCCAGATGCCCCCTCTCTGCT 1796  
QY 722 GATCTTCAGGATTTCTTGAGTGGATCTTTCAAGTACCCCGGCGCAGAGTCCCAATGGG 781  
Db 1797 GACCTTCAGGATTTCTTGAGCGGATCTTTCAAGTACCCCGGCGCAGAGTGTCCACGGG 1856  
QY 782 AACTCTTTGAGCTCTCTCAGCCTGCCCCCTGGAGCGCTGCAGCCTCTTAAGCCCAACAGC 841  
Db 1857 AACTCTTTGAGCTCTCTCAGCCTGCCCCCTGGGCGCCTGCGAGCCTTCCAAGCCCAACAGC 1916  
QY 842 ACAGTACCCCAAGGAGAACCCAAACCTAAGCGCGCGAAGAAAGTGGAGGCGCTTCCAA 901  
Db 1917 ACAGTACCCCAAGGAGAACCCAAACCTAAGCGCGCGAAGAAAGTGGAGGCGCTTCCAA 1976  
QY 902 CGTTGATGCCCCCTCTCTCTTCTTCAATCAATGTGAGGAGTCAAAAGGCGCTGT 955  
Db 1977 CGTTGACACCCCTCTCTCTTCTTCTTCAATCAATGTGAGGAGTCAAAAGGCGCTGT 2030  
RESULT 11  
BC011146  
LOCUS  
DEFINITION Mus musculus, RIKEN cDNA 5730551F12 gene, clone MGC:19282  
IMAGE:4016209, mRNA, complete cds.  
ACCESSION BC011146  
VERSION BC011146.1 GI:15029845  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2402)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (25-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LILNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 24 Row: a Column: 19

ACCESSION	AY027881	AY027881.1	GI:15777192
VERSION			
KEYWORDS			
SOURCE	Bos taurus	(cow)	
ORGANISM	Bos taurus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.		
REFERENCE	1 (bases 1 to 3159)		
AUTHORS	Rinck, G., Birghan, C., Hatada, T., Meyers, G., Thiel, H. J. and Tautz, N.		
TITLE	A cellular J-domain protein modulates polyprotein processing and cytopathogenicity of a pestivirus		
JOURNAL	J. Virol. 75 (19), 9470-9482 (2001)		
MEDLINE	21424530		
PUBMED	11533209		
REFERENCE	2 (bases 1 to 3159)		
AUTHORS	Rinck, G., Tautz, N. and Meyers, G.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-FEB-2001) Institut fuer Virologie (FB10), Justus-Liebig-Universitaet Giessen, Frankfurterstrasse 107, Giessen 35392, Germany		
FEATURES	Location/Qualifiers		
source	1..3159		
	/organism="Bos taurus"		
	/mol_type="mRNA"		
5'UTR	/db_xref="taxon:9913"		
CDS	1..187		
	188..2287		
	/function="interacts with nonstructural protein 2 of bovine viral diarrhoea virus"		
	/notes="contains putative zinc finger domain"		
	/codon_start=1		
	/product="J-domain protein Jiv"		
	/protein_id="AAK28650.1"		
	/db_xref="GI:15777193"		
	/translation="MAQKIPGEGSLCAHHSGGASLRTLPSPVDPDILSFGSLRDSAG SAPNGRLCTEHSSKYTPPPNAHSDSPHGPPLAEEDPDQSEASGESEGV DQLSRENTGYQDQNSGFLSIPSTCNQGTGPIPEGPYSEGRDSSSNFCHHCTSP ALGEDLEGEYDEEPLKVPSPSVPEKKPAPRQRHVRPAKEDTREGGRDRPS PQRHRLGRKSQADKRGVGLMGAEELCOLGOAGFWLLLELLVLGVEYVETCGHLLYA CRLQKSDLDLLAVWVGVWAGRLRWQAQVWFQSLSGFCYGAGLGFTRFLRLVALLLL ALIALGLDELFWVGLUSDRUGRDKATWIFSLASPTWQRCLLLRESRPQQLV RVQWGWLELPMWKYQRTNAQAPVAGRGYCOPEEVARLLTMAGVPEDELNPFHVLG VITQADDELPAQYLAQYLAVMVHPDKNHHFRAEAFKVLRAAMDIVSNPERRKEYEMKR MAENELSRVNEFLSKLQAMNTMCSRCQGGKRRHREMPREPKSARYCAECNRLHPAE EGFDAWESSMLGIKITYFALMDGKVYDITWAGACQGVSPDTHRVPIHISGSRNPFG TSGROPATDPAADLQDLSLRIFQVPPQGMNGNFFAAPQFGFGATAASKPNSTVPK GAQKPRKRRKPRFOR"		
3'UTR	2288..3159		
	/notes="use of alternative polyadenylation site results in 3'UTR longer than in the sequence presented in GenBank"		
BASE COUNT	755 a 759 c 895 g 750 t		
ORIGIN			
Query Match	72.7%;	Score 834;	DB 4; Length 3159;
Best Local Similarity	92.3%;	Pred. No. 2e-258;	
Matches	906;	Conservative	0; Mismatches 60; Indels 16; Gaps 2;
QY	2	CCTTGGGTCAAGCAGAAATATTATAGCGAGGGGAATGCACCTTAGCTAGTGGGCGCTAC	61
Db	1394	CCTTGGGTCAACAGAGAGCAATAGGCAGGCGAATGCACCTTAGCTAGTGGTCTGTTAC	1453
QY	62	TGCCAGCTGAGAGAACTGCTGCACCTTGACCATCGCTGGGGTTCCTGAGGATGAG	121
Db	1454	TGCCAGCTGAGAGAAAGTGGCTCGACCTTGACCATCGCTGGGGTTCCTGAGGATGAG	1513
QY	122	CTAAACCCCTTTCCATCTACTCTGGGGTGTGAGGCCACAGCATCAGATCTTCAACTGAAGAAG	181
Db	1514	CTAAACCCCTTTTACGTGTGGGGTGTGAGGCCACAGCATCAGATCTTGAACCTGAAGAAG	1573
QY	182	GCCTATAGACGCTGCGACGTGATGGTTCACTCTGACAAAAATCATCATCCCGGGCTGAG	241

QY 1042 GCTCTGTTGCCAGCTGGAGTGGTGTGATCTCAGCTTACTGCAACTCTGTCTCTCC 1101  
Db 781 GCTCTGTTGACAGCTGGAGTGGTGTGATCTCAGCTTACTGCAACTCTGTCTCTCC 840  
QY 1102 CGGTTTCAAGCAATTCCTCCATCTCAGCCTCTCAGTAGCTGGGAT 1147  
Db 841 CGGTTTCAAGCAATTCCTCCATCTCAGCCTCTCAGTAGCTGGGAT 886

RESULT 8  
AY027882 2402 bp mRNA linear MAM 25-SEP-2001  
LOCUS Bos taurus J-domain protein Jiv mRNA, complete cds.  
DEFINITION AY027882  
ACCESSION AY027882.1 GI:15777194  
VERSION  
KEYWORDS  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 2402)  
AUTHORS Rinck, G., Birghan, C., Hatada, T., Meyers, G., Thiel, H. J. and Tautz, N.  
TITLE A cellular J-domain protein modulates polyprotein processing and  
cytopathogenicity of a pestivirus  
JOURNAL J. Virol. 75 (19), 9470-9482 (2001)  
MEDLINE 21424530  
PUBMED 11533209  
REFERENCE 2 (bases 1 to 2402)  
AUTHORS Rinck, G. and Tautz, N.  
TITLE Direct Submission  
JOURNAL Submitted (22-FEB-2001) Institut fuer Virologie (FB10),  
Justus-Liebig-Universitaet Giessen, Frankfurterstrasse 107, Giessen  
35392, Germany

FEATURES  
source 1..2402  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/cell\_line="MDBK"  
1..187  
188..2287  
/functions="interacts with nonstructural protein 2 of  
bovine viral diarrhoea virus"  
/notes="contains putative zinc finger domain"  
/codon\_start=1  
/product="J-domain protein Jiv"  
/protein\_id="AAK28651.1"  
/db\_xref="GI:15777195"  
/translation="MAQKHGEGGLGCAHSHGASLRTLGPSVDPDILSPSGLRDSAG  
SAPNGTRCLTEHSPKYPNPAHWSDPHGPGRGPLAPEDPDQSEASFEESGV  
DOELRENETGYDDGNSFLSPSTCNCGTPIEPGYSSEGRDSSSNFCHCTSP  
ALGDEBELGEYDEEPLKPPSDVSRVPSEKKPAPRQRHVRPAKEDTREGGRDPRS  
PGRHRLGRKRSQADKRGRLGWABELCQLGQAFWMLIELLVLVGEYVETGHLIYA  
CQLKGSDDLRLVWVWAGRLGWQVNFQLSQFCYAGLFRFLRLVGLALLL  
ALALLGLCQLGWRFLVGLSDRLGWRDKATWIFSLASPTWQRLILLRESRPWQOLV  
RIVQWGLELPWQRTNQAAPVAGRYCQPEBEVARILLTMAGVPEDELNPFHVLG  
VEATASDVELKAYROLAVMHPDKNHPRAEAFVKVLAANDIVSNPERRKEYEWR  
MENEISRSVNEFLSKLOAMNTMCSRCQGRHREMDREPKSARYCAEONRLHPAE  
EGDFWAESSMLGUKITIFALMDGKYDITETWAGCQVRGSPDTHRYVPHISFGSRMPG  
TSGQRATPDAPADLQDFLSRIFQVPPGQMSNGNFFAAPQPGPGATAASKPNSTVPK  
GEAKPRKRRKVRPFQR"  
2288..2402  
/note="use of alternative polyadenylation site results in  
3'UTR shorter than in the sequence presented in GenBank  
Accession Number AY027881"  
BASE COUNT 559 a 582 c 734 g 527 t  
ORIGIN

Query Match 72.7%; Score 834; DB 4; Length 2402;  
Best Local Similarity 92.3%; Pred. No. 2e-258;  
Matches 906; Conservative 0; Mismatches 60; Indels 16; Gaps 2;

QY 2 CTTTGGGTCAAGCAGATATTAATAGCAGGGGAATGCACCTGTAGTCTAGTGGCGCTAC 61  
Db 1394 CTTTGGGTCAAAACAGAGGACCAATAGCAGCGCAATGCACCTGTAGTCTAGTGGTTCGTTAC 1453  
QY 62 TGCAGGCTCAAGAGGAAGTGGCTCGACTCTTGCACCATGCTGGGGTTCCTGAGGATGAG 121  
Db 1454 TGCAGGCTCAAGAGGAAGTGGCTCGACTCTTGCACCATGCTGGGGTTCCTGAGGATGAG 1513  
QY 122 CTAAACCCCTTTCCATGCTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACGAAGAAG 181  
Db 1514 CTAAACCCCTTTCCATGCTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACGAAGAAG 1573  
QY 182 GCCTATAGCAGCTGGCAGTGTGTTTCATCTCTGACAAAAATCATCATCCCGGGCTGAG 241  
Db 1574 GCCTATAGCAGCTGGCAGTGTGTTTCATCTCTGACAAAAATCATCATCCCGGGCTGAG 1633  
QY 242 GAGGCTTCAAGCTTTTGCAGCAGCTTGGGACATTTGTACCAATGCTGAAGACGAAG 301  
Db 1634 GAGGCTTCAAGCTTTTGCAGCAGCTTGGGACATTTGTACCAATGCTGAAGACGAAG 1693  
QY 302 GAGTATGAGATGAAGCAATGGCAGAGAATGAGCTGAGCCGGTCAGTAAATGAGTTTCTG 361  
Db 1694 GAATATGAGATGAAGCAATGGCAGAGAATGAGCTGAGCCGGTCAGTAAATGAGTTTCTG 1753  
QY 362 TCCAAAGCTCAAGATGACCTTCAAGGAGCAATGAATCTATGATGTTAGCCGATGCCAA 421  
Db 1754 TCCAAAGCTGCAA-----GAAGCAATGAATGATGATGTCACCCGATGCCAG 1801  
QY 422 GGAAGCATAGGAGGTTTGAATGAGCCGGGAACCTTAAGAGTGCAGATATCTGTCTGAG 481  
Db 1802 GGAAGCATAGGAGGTTTGAATGAGCCGGGAACCTTAAGAGTGCAGATATCTGTCTGAG 1861  
QY 482 TGTAAATAGCTGCATCTGCTGAGGAGGAGAGCTTTTGGCAGAGTCAAGCATGTTGGGC 541  
Db 1862 TGTAAATAGCTGCATCTGCTGAGGAGGAGTGTCTTTTGGCAGAGTCAAGCATGTTGGGC 1921  
QY 542 CTCAGATCACCTACTTTGCACCTGATGGATGGAAGGTGTATGACATACAGAGTGGGCT 601  
Db 1922 CTCAGATCACCTACTTTGCACCTGATGGATGGAAGGTGTATGATATCAGAGTGGGCT 1981  
QY 602 GGATGCCAGGTGTAGGTATCTCCAGATACCCACAGAGTCCCTTATCATCTCATTT 661  
Db 1982 GGATGCCAGGTGTAGGTATCTCCAGATACCCACAGAGTCCCTTATCATCTCATTT 2041  
QY 662 GGTTCCTGGATTCACAGGACAGAGGCGGCGAGAGCCACCCAGATGCCCTCTCTGCT 721  
Db 2042 GGTTCCTGGATTCACAGGACAGAGGCGGCGAGAGGACTACTCCAGATGCCCTCTCTGCT 2101  
QY 722 GATCTTTCAGGATTTCTTGATCGGATCTTTCAAGTACCCCGGCGAGATGCCCAATGG 781  
Db 2102 GATCTTTCAGGATTTCTTGATCGGATCTTTCAAGTACCCCGGCGAGATGTCCAACGG 2161  
QY 782 AACTTCTTTGCAGCTCCTCAGCCTGCGCTGGAGCGCGCTGCAGCCTCTAAGCCCAACAGC 841  
Db 2162 AACTTCTTTGCAGCTCCTCAGCCTGCGCTGGAGCGCGCTGCAGCCTCTAAGCCCAACAGC 2221  
QY 842 ACAGTACCAAGGAGAACCCAAACCTAAGCGCGGAGAAAGTGAAGGAGGCCCTTCCAA 901  
Db 2222 ACAGTACCAAGGAGAACCCAAACCTAAGCGCGGAGAAAGTGAAGGAGGCCCTTCCAA 2281  
QY 902 CGTTGATGCCCTTCTCTTCTCTCAATCAATGTGAGGAGTCAAAAGGGCTGT----AG 957  
Db 2282 CGTTGATGCCCTTCTCTTCTCTCAATCAATGTGAGGAGTCAAAAGGGCTGT----AG 957  
QY 958 CACAGGATGGAGTTTGATTTAT 979  
Db 2342 CACCAGATGGAGTTTGATTTGT 2363

RESULT 9  
AY027881 3159 bp mRNA linear MAM 25-SEP-2001  
LOCUS Bos taurus J-domain protein Jiv mRNA, complete cds.  
DEFINITION

Db	271	GGCTTAGACAGCTGCGCATGTGTTTCATCTCTGACAAATATCATATCCCCGGGCTGA	330
Qy	241	GGAGGCGCTTCAAGGCTTTTGGAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAA	300
Db	331	GGAGGCGCTTCAAGGCTTTTGGAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAA	390
Qy	301	GGAGTATGAGATGAACGAATGCGAGAAATGAGTGGCCGTCAGTAAATGAGTTTCT	360
Db	391	GGAGTATGAGATGAACGAATGCGAGAAATGAGTGGCCGTCAGTAAATGAGTTTCT	450
Qy	361	GTCCAGCTGCAAGTACCTCAAGGCGCAATGAATACTATGATGTGTAGCGATGCCA	420
Db	451	GTCCAGCTGCAAGTACCTCAAGGCGCAATGAATACTATGATGTGTAGCGATGCCA	510
Qy	421	AGGAAGCATAGGAGGTTTCAATGAGCCGCGCAATGAAGTGCAGATACTGTGCTGA	480
Db	511	AGGAAGCATAGGAGGTTTCAATGAGCCGCGCAATGAAGTGCAGATACTGTGCTGA	570
Qy	481	GTGTAATAGCTGCATCTCTCTGAGGAAGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	540
Db	571	GTGTAATAGCTGCATCTCTCTGAGGAAGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	630
Qy	541	CCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTATGACATCAGAGTGGGC	600
Db	631	CCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTATGACATCAGAGTGGGC	690
Qy	601	TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATCACATCTCAT	660
Db	691	TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATCACATCTCAT	750
Qy	661	TGGTTCTCGGATTCAGGACCCAGAGGCGCGAGAGCCACCCAGATCCCTCTCTGC	720
Db	751	TGGTTCTCGGATTCAGGACCCAGAGGCGCGAGAGCCACCCAGATCCCTCTCTGC	810
Qy	721	TGATCTTCAGGATTTCTTGTAGTCCGATCTTTTCAAGTACCCCGGCGAGTGCCTCAATGG	780
Db	811	TGATCTTCAGGATTTCTTGTAGTCCGATCTTTTCAAGTACCCCGGCGAGTGCCTCAATGG	870
Qy	781	GAATCTTTTGAGCTCTCTCAGCTGCCCTGAGCGCTGAGCGCTCTTAAGCCCAACAG	840
Db	871	GAATCTTTTGAGCTCTCTCAGCTGCCCTGAGCGCTGAGCGCTCTTAAGCCCAACAG	930
Qy	841	CACAGTACCCAGGAGAGCCCAACCTTAAGCGCGGAGAAAGT	885
Db	931	CACAGTACCCAGGAGAGCCCAACCTTAAGCGCGGAGAAAGT	975
RESULT 7			
AF351784			
LOCUS	AF351784 953 bp mRNA linear PRI 24-MAY-2001		
DEFINITION	Homo sapiens dopamine receptor interacting protein mRNA, partial cds.		
ACCESSION	AF351784		
VERSION	AF351784.1 GI:14194056		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 953)		
TITLE	Bermak, J.C., Li, M., Bullock, C. and Zhou, Q.Y.		
JOURNAL	Regulation of transport of the dopamine D1 receptor by a new		
MEDLINE	membrane-associated ER protein		
PUBMED	Nat. Cell Biol. 3 (5), 492-498 (2001)		
REFERENCE	11331877		
AUTHORS	2 (bases 1 to 953)		
TITLE	Bermak, J.C., Li, M., Bullock, C.M. and Zhou, Q.-Y.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (21-PEB-2001) Pharmacology, University of California,		
PUBMED	19182 Jamboree Blvd., Irvine, CA 92697, USA		
FEATURES	Location/Qualifiers		
source	1. .953		

/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="12"			
<1..644			
/notes="DRIP78; ER-membrane protein; zinc finger motif"			
/codon_start=3			
/product="dopamine receptor interacting protein"			
/protein_id="AAK56241.1"			
/db_xref="GI:14194057"			
/translation="ANDIVSNABKRYEYEMKMAENELSRVNEFLUSKLODDLEKAMN			
TMMSCRCQKHRRFENDREPKSARYCAECNRLHPABEGDFWAESSMLGLKITFYFALMD			
KGYVDITWAGCQGVISPDTHRPVPHISFGSRIPGTRGRATPDAPPADLODFLSR			
IFQVPPQMPNAGNFFAPOPAPGAAAAASKPNSTVPKGEAKPKRKKVRRPQR"			
BASE COUNT	240 a	236 c	246 g
ORIGIN	231 t		
Query Match 75.6%; Score 867.2; DB 9; Length 953;			
Best Local Similarity 99.4%; Pred. No. 3.1e-269;			
Matches 881; Conservative 0; Mismatches 3; Indels 2; Gaps 1;			
Qy	264	CAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAAGGAGTATGAGATGAAACGAATGG	323
Db	1	CAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAAGGAGTATGAGATGAAACGAATGG	60
Qy	324	CAGAGAATGAGCTGAGCGGTCAAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGACCTCA	383
Db	61	CAGAGAATGAGCTGAGCGGTCAAGTAAATGAGTTTCTGTCCAAGCTGCAAGATGACCTCA	120
Qy	384	AGAGGCAATGAATATCTATGATGTAGCGGATGCCAAGGAAAGCATAGAGGTTTGAAA	443
Db	121	AGGAGGCAATGAATATCTATGATGTAGCGGATGCCAAGGAAAGCATAGAGGTTTGAAA	180
Qy	444	TGGACCGGGAACCTAAGAGTGGCAGATAGTGTCTGAGTGTAAATAGCTGCATCTCTGCTG	503
Db	181	TGGACCGGGAACCTAAGAGTGGCAGATAGTGTCTGAGTGTAAATAGCTGCATCTCTGCTG	240
Qy	504	AGGAAGGAGACTTTTGGGCGAGCTCAAGCATGTTGGGCGCTCAAGATCACTACTTTGCAC	563
Db	241	AGGAAGGAGACTTTTGGGCGAGCTCAAGCATGTTGGGCGCTCAAGATCACTACTTTGCAC	300
Qy	564	TGATGGATGGAAGGTGTATGACATCACAGAGTGGGCTGGATGCCAGCGTGTAGGTATCT	623
Db	301	TGATGGATGGAAGGTGTATGACATCACAGAGTGGGCTGGATGCCAGCGTGTAGGTATCT	360
Qy	624	CCCCAGATACCCAGAGTCCCCCTATCACATCTCATTTGGTTCTCGGATTCAGGCACCA	683
Db	361	CCCCAGATACCCAGAGTCCCCCTATCACATCTCATTTGGTTCTCGGATTCAGGCACCA	420
Qy	684	GAGGCGGCGAGAGAGCCACCCAGATGCCCCCTCTCTGATCTTCAGGATTTCTTGAGTC	743
Db	421	GAGGCGGCGAGAGAGCCACCCAGATGCCCCCTCTCTGATCTTCAGGATTTCTTGAGTC	480
Qy	744	GGATCTTTCAAGTACCCCGAGGCGAGATGCCAATGGGAACTTTCTTTGCAGCTCCTCAGC	803
Db	481	GGATCTTTCAAGTACCCCGAGGCGAGATGCCAATGGGAACTTTCTTTGCAGCTCCTCAGC	540
Qy	804	CTGCCCTTGAGCGCGCTGAGCTCTTAAGCCCAACAGCAGTACCCAGGGAGAGGCCA	863
Db	541	CTGCCCTTGAGCGCGCTGAGCTCTTAAGCCCAACAGCAGTACCCAGGGAGAGGCCA	600
Qy	864	AACCTAAGCGGCGGAAGAAAGTGAAGAGGCGCTTCCAAAGTTCATGCCCTTCTCTTCC	923
Db	601	AACCTAAGCGGCGGAAGAAAGTGAAGAGGCGCTTCCAAAGTTCATGCCCTTCTCTTCC	660
Qy	924	TCAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAGGATGGAGTTTGAATATCCCT	983
Db	661	TCAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAGGATGGAGTTTGAATATCCCT	720
Qy	984	CCTCCCCCAACCTAGGAACCTGAATCTTTTTT - CTTTTATTTTTTGAGATGGAGTCTT	1041
Db	721	CCTCCCCCAACCTAGGAACCTGAATCTTTTTTCTTTTTTAAATTTTTTGTAGATGGAGTCTT	780

QY	241	GGAGGCTTCAAGGTTTGGAGCAGCTTGGACATTTGACCAATGCTGAAAGCGAAA	300
Db	117982	GGAGGCTTCAAGGTTTGGCGCAGTTGGGACATTTGACCAATGCTGAAAGCGAAA	118041
QY	301	GGAGTATGAGATGAAACGAATGGCAGAGAAATGAGTGGCGGTCAGTAAATGAGTTCT	360
Db	118042	GGATATGAGATGAAAGGACGGCAGAGATCAGCTGAGCGGTCAGTAAATGAGTTCT	118101
QY	361	GTCGAAGTGAAGATGACCTCAAGGAGGCAATGAATCTATGATGTGTAGCCGATGCCA	420
Db	118102	GTCCAAGCTGCAAGATGACCTCAAGGAGGCGAGTGAATCTGATGTGTGCGGATGCCA	118161
QY	421	AGGAAGCATAGGAGTTTGAATTCGACCGGCACTTAAGTGCACAGATCTGTGCTGA	480
Db	118162	AGGAAGCATAGGAGTTTGAAGTTCGACCGGCACTTAAGTGCACAGATCTGTGCTGC	118221
QY	481	GTGTAATAGGCTGATCTCTGTGAGGAGGAGACTTTTGGCAGAGTCAAGCATGTTGGG	540
Db	118222	GTGTATAGGCTGCTCTGCTGAGGAGGAGACTTTTGGCAGAGTCAAGCATGTTGGG	118281
QY	541	CCTCAAGATCACCTACTTTGCACCTGATGGATGGAAGGTGTATGACATCAGAGTGGGC	600
Db	118282	CCTCAAGATCACCTACTTTAGCACTGATGGATGGAAGGTGTATGACACACACAGTGGGC	118341
QY	601	TGATGCGCAGGTGATGATCTCCCGAGATACCCACAGAGTCCCTTATCAATCTCAT	660
Db	118342	TGAGGCGCAGGTCTCTAG-ATATCTCCCGAGACCCACAGAGTCTCTCATCTCTCAT	118400
QY	661	TGTTCTCGGATTCAGGACCCAGAGGCGGCGAGAGGACCCACCCAGATCCCTCTGTC	720
Db	118401	TGTTCTCGGATTCAGGACCCAGAGGCGGCGAGAGGACCCACCCAGATCCCTCTGTC	118460
QY	721	TGATCTTCAGGATTTCTGAGTGGATCTTTCAAGTACCCCGGAGGAGTCCCAATGG	780
Db	118461	TGATCTTCAGGATTTCTGAGTGGATCTTTCAAGTACCCCGGAGGAGTCCCAATGG	118520
QY	781	GAATCTTTTGCAGTCTCTCAGGCTGCGCTGGAGCGCTGCGAGCTCTTAAGCCCAACAG	840
Db	118521	GAATCTTTTGCAGTCTCTCAGGCTGCGCTGGAGCGCTGCGAGCTCTTAAGCCCAACCG	118580
QY	841	CACAGTACCCAGGAGAACCCAACTAGAGCGCGGAGAAAGTGAAGGCGCTTCCA	900
Db	118581	CACAGTACCCAGGAGAAACGAAACCGAAGCGCGGAGAAAGTGAAGGCGCTTCCA	118640
QY	901	ACGTTG-ATGCGCTCTCTCTTCTCTCAATGTCAGGAGTCAAAAGGCTGT----	955
Db	118641	ACGTTGAAGCCCTCTCTCTTCTCTCAAAATGTCAGGAGTGAAGGAGTGTGTAC	118700
QY	956	AGCAGAGGATGAGTTTGAATTTATCCCTCTCTCCCGCAACACCTAGGAACCTGAT	1015
Db	118701	AGCGCGGATGAGTTTGAATTTCTCTCTCTCTCCCGCAACACCTAGGAACCTGAT	118760
QY	1016	CTTTTATTTTGAATGAGTCTCTGCTGTGTTGCCAGTGGAGTGGTGTGTAT	1075
Db	118761	-TTTTTTTTTTGAGACAGAGTCTGCTGTGTTGCCAGTGGAGTGGTGTGTAT	118819
QY	1076	CTCAGCTTACTGCAACTCTGCTCCCGGTTCAAGCAATTTCCCATCTCAGCTCTCTG	1135
Db	118820	CTTGGTTTACTGCAACTCTGCTCTGCTGTTTCAAGCAATTTCCCATCTCAGCTCTCTG	118879
QY	1136	AGTAGCTGGAT 1147	
Db	118880	AGTAGCTGGAT 118891	
RESULT 6			
BC016941			
LOCUS			
DEFINITION Homo sapiens, clone MGC:21452 IMAGE:3448446, mRNA, complete cds.			
ACCESSION BC016941			
VERSION BC016941.1 GI:16877383			
KEYWORDS MGC.			

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1818)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
REMARK	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Location/Qualifiers
source	1..1818 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:21452 IMAGE:3448446" /tissue_type="Placenta, choriocarcinoma" /clone_id="NIH MGC_10" /lab_host="DH105" /note="Vector: pCMV-SPORT6" 401..1570 /codon_start=1 /product="Unknown (protein for MGC:21452)" /protein_id="AAH16941.1" /db_xref="GI:16877384" /translation="MKRMAENLSRVNEFLSKLQDDL'EAMNTMWCSCQGHRRFE MDREPKSARYCAECNLRHPABEGDFWABSSMLGLKITYFALMDGVYDITWAGCCORV GTSPTHRVYHISFGSRIPGTRQAPADLPADLOFLSRIFOVPGWPNQFV AAOPOAGAAASKPNSTVPKGEAKPKRKLAEKQELARGLETKGIKQDLIHLRQ AYLESHABEENEDVLGDTEETETKPIELPVKEEPEPKTVDVAEEKVKITSEI PTERNQKAEFRNPVSLSKVAAARFGISSVPTKGLSDNPMVNLDKKERAQ REGLNVSISRSKDEKLKKRERFGIVTSAGTGTEDTEAKKRKRAERFGIA"
CDS	BASE COUNT 532 a 373 c 494 g 419 t ORIGIN
Query Match	77.0%; Score 883.4; DB 9; Length 1818;
Best Local Similarity	99.9%; Pred. No. 1.9e-274;
Matches	884; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 GCCTTGGGTCACAGAGAATATTAAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 60
Db	91 GCCTTGGGTCACAGAGAATATTAAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 150
QY	61 CTGCCAGCCTGAAGAGGAAGTGGCTCGACTCTTGACCATGGCTGGGTTCTCTGAGGATGA 120
Db	151 CTGCCAGCCTGAAGAGGAAGTGGCTCGACTCTTGACCATGGCTGGGTTCTCTGAGGATGA 210
QY	121 GCTAAACCTTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACGAAGA 180
Db	211 GCTAAACCTTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACGAAGA 270
QY	181 GGCCTATAGACAGCTGGCAGTGATGGTTTCATCTTGACAAAATCATATCCCGGCTGA 240

QY 121 GCTAAACCCCTTTTCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACTGAAGAA 180  
Db 1415 GCTAAACCCCTTTTCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACTGAAGAA 1474  
QY 181 GGCTATAGACAGTGCAGTGTGTTCTCTGACAAAATCATCATCCCGGGCTGA 240  
Db 1475 GGCTATAGACAGTGCAGTGTGTTCTCTGACAAAATCATCATCCCGGGCTGA 1534  
QY 241 GGAGGCCCTTCAAGGTTTGGCAGCAGCTTGGGACATTGTGAGCAATGCTGAAAAGCGAAA 300  
Db 1535 GGAGGCCCTTCAAGGTTTGGCAGCAGCTTGGGACATTGTGAGCAATGCTGAAAAGCGAAA 1594  
QY 301 GGAGTATGAGATGAACGAATGGCAGAGATGAGTCTAGCCGGTCAAGTAATGAGTTTCT 360  
Db 1595 GGAGTATGAGATGAACGAATGGCAGAGATGAGTCTAGCCGGTCAAGTAATGAGTTTCT 1654  
QY 361 GTCACAGCTCAAGATGACCTCAAGGAGCAATGAATCTATGATGTGACCGATGCCA 420  
Db 1655 GTCACAGCTCAAGATGACCTCAAGGAGCAATGAATCTATGATGTGACCGATGCCA 1714  
QY 421 AGAAAAGCATAGGAGGTTTGAATGGACCGGAACTTAAGAGTGCCAGATACTGTGCTGA 480  
Db 1715 AGAAAAGCATAGGAGGTTTGAATGGACCGGAACTTAAGAGTGCCAGATACTGTGCTGA 1774  
QY 481 GTCTAATAGCTGATCTCTGCTGAGGAAGGAGACTTTTGGCCAGAGTCAAGCATGTTGGG 540  
Db 1775 GTCTAATAGCTGATCTCTGCTGAGGAAGGAGACTTTTGGCCAGAGTCAAGCATGTTGGG 1834  
QY 541 CTTCAAGATCACCTACTTTGCATCTGATGGATGGAAGGTGTATGACATCAAGAGTGGC 600  
Db 1835 CTTCAAGATCACCTACTTTGCATCTGATGGATGGAAGGTGTATGACATCAAGAGTGGC 1894  
QY 601 TGGATGCCAGCGTGTAGTATCTCCAGATACCCAGAGTCCCTATCACATCTCATT 660  
Db 1895 TGGATGCCAGCGTGTAGTATCTCCAGATACCCAGAGTCCCTATCACATCTCATT 1954  
QY 661 TGGTTCTCGGATTTCCAGGCCACAGAGGCGGCGAGAGCCACCCAGATCCCTCTCTGC 720  
Db 1955 TGGTTCTCGGATTTCCAGGCCACAGAGGCGGCGAGAGCCACCCAGATCCCTCTCTGC 2014  
QY 721 TGATCTCAGGATTTCTTGATCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGG 780  
Db 2015 TGATCTCAGGATTTCTTGATCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGG 2074  
QY 781 GAATCTTTTGAGTCTCTCAGCTGCCCTGAGCGGCTGAGCGCTTAAGCCCAACAG 840  
Db 2075 GAATCTTTTGAGTCTCTCAGCTGCCCTGAGCGGCTGAGCGCTTAAGCCCAACAG 2134  
QY 841 CACAGTACCCAGGGAGAACCCAAACCTAAGCGCGGAGAAAGTGAAGGCGCTTCCA 900  
Db 2135 CACAGTACCCAGGGAGAACCCAAACCTAAGCGCGGAGAAAGTGAAGGCGCTTCCA 2194  
QY 901 ACCTTGATGCCCTTCTCTTCTCAAAATCAATGTGAGGAGTCAAAAGGCGTGTAGCAC 960  
Db 2195 ACCTTGATGCCCTTCTCTTCTCAAAATCAATGTGAGGAGTCAAAAGGCGTGTAGCAC 2254  
QY 961 AGATGGAGTTGATTTATCTCTCTCCCAACACCTAGGAACCTGAATCTTTTCTTTT 1020  
Db 2255 AGATGGAGTTGATTTATCTCTCTCCCAACACCTAGGAACCTGAATCTTTTCTTTT 2314  
QY 1021 TATTTTTCAGATGGAGTCTGCTCTGTCGCCAGCTGGAGTGCAGTGTGATCTCAG 1080  
Db 2315 TATTTTTCAGATGGAGTCTGCTCTGTCGCCAGCTGGAGTGCAGTGTGATCTCAG 2374  
QY 1081 CTTACTGCAACCTCTGCTCTCCCGGTTCAAGCAATTTCTCCACTCTCAGCTCTCTGAGTAG 1140  
Db 2375 CTTACTGCAACCTCTGCTCTCCCGGTTCAAGCAATTTCTCCACTCTCAGCTCTCTGAGTAG 2434  
QY 1141 CTGGGAT 1147  
Db 2435 CTGGGAT 2441

RESULT 5  
AC129837 190201 bp DNA linear HTG 11-JUN-2003  
LOCUS Papio anubis clone rp41-88j2, WORKING DRAFT SEQUENCE.  
DEFINITION AC129837  
ACCESSION AC129837  
VERSION AC129837.16 GI:30725963  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
SOURCE Papio anubis (olive baboon)  
ORGANISM Papio anubis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
Cercopitheciinae; Papio.  
REFERENCE 1 (bases 1 to 190201)  
AUTHORS Prescott,A., Shaikh,T. and Roe,B.A.  
TITLE Papio anubis BAC Clone rp41-88j2  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 190201)  
AUTHORS Prescott,A. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2002) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
REFERENCE 3 (bases 1 to 190201)  
AUTHORS Prescott,A., Shaikh,T. and Roe,B.A.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,  
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,  
OK 73019, USA  
COMMENT OK May 15, 2003 this sequence version replaced gi:30698658.  
----- Genome Center  
Center: Department Of Chemistry And Biochemistry  
The University Of Oklahoma  
Center code:UOKNOR  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 190201: contig of 190201 bp in length.  
Location/Qualifiers  
1..190201  
/organism="Papio anubis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9555"  
/clone="rp41-88j2"  
/clone\_lib="RPCI - 41 Male (Olive) Baboon BAC Library"  
BASE COUNT 46020 a 50800 c 49199 g 44180 t 2 others  
ORIGIN

Query Match 84.0%; Score 963.6; DB 2; Length 190201;  
Best Local Similarity 92.1%; Pred. No. 2.9e-300;  
Matches 1061; Conservative 0; Mismatches 84; Indels 7; Gaps 4;  
QY 1 GCCTTGGGTCAAGAGGAAGTGGCTCGACTTTGACCATGGGTGGGGTTCCTGAGGATGA 117801  
Db 11742 GCCTTGGGTCAAGAGGAAGTGGCTCGACTTTGACCATGGGTGGGGTTCCTGAGGATGA 117801  
QY 61 CTGCCAGCTCAAGAGGAAGTGGCTCGACTTTGACCATGGGTGGGGTTCCTGAGGATGA 120  
Db 117802 CTGCCAGCTCAAGAGGAAGTGGCTCGACTTTGACCATGGGTGGGGTTCCTGAGGATGA 117861  
QY 121 GCTAAACCCCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACTGAAGAA 180  
Db 117862 GCTAAACCCCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACTGAAGAA 117921  
QY 181 GGCCTATAGACAGTGGCAGTGTGTTCTCTGACAAAATCATCATCCCGGGCTGA 240  
Db 117922 GGCCTATAGACAGTGGCAGTGTGTTCTCTGACAAAATCATCATCCCGGGCTGA 117981

QY	241	GGAGGCTTCAAGGTTTGGCAGCAGCTTGGGACATTTGTACGAATGCTGAAAGCGAAA	300
Db	685	GGAGGCTTCAAGGTTTGGCAGCAGCTTGGGACATTTGTACGAATGCTGAAAGCGAAA	744
QY	301	GGAGTATGAGATGAAACGAATGGCAGAGATGAGTCAGCCGGTCAGTAAATGAGTTTCT	360
Db	745	GGAGTATGAGATGAAACGAATGGCAGAGATGAGTCAGCCGGTCAGTAAATGAGTTTCT	804
QY	361	GTCCAAAGCTCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTAGCCGATGCCA	420
Db	805	GTCCAAAGCTCAAGATGACCTCAAGGAGGCAATGAATACTATGATGTAGCCGATGCCA	864
QY	421	AGGAAGCATAGGAGTTTGAATGGACCGGAACTTAAGAGTGCAGATACTGTGCTGA	480
Db	865	AGGAAGCATAGGAGTTTGAATGGACCGGAACTTAAGAGTGCAGATACTGTGCTGA	924
QY	481	GTGTAATAGGCTGATCCTGCTGAGGAGGAGACTTTTGGCAGAGTCAAGCATGTTGGG	540
Db	925	GTGTAATAGGCTGATCCTGCTGAGGAGGAGACTTTTGGCAGAGTCAAGCATGTTGGG	984
QY	541	CCTCAAGATCACCTACTTTTGCACTGATGGATGGAAGGTGTATGACATCACAGAGTGGC	600
Db	985	CCTCAAGATCACCTACTTTTGCACTGATGGATGGAAGGTGTATGACATCACAGAGTGGC	1044
QY	601	TGATGCGAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATACATCTCAT	660
Db	1045	TGATGCGAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATACATCTCAT	1104
QY	661	TGGTCTCGGATTCAGGACACAGGCGCGCAGAGGACACCCAGATGCCCTCTCTGC	720
Db	1105	TGGTCTCGGATTCAGGACACAGGCGCGCAGAGGACACCCAGATGCCCTCTCTGC	1164
QY	721	TGATCTTCAGGATTTCTTGTAGTCGGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG	780
Db	1165	TGATCTTCAGGATTTCTTGTAGTCGGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG	1224
QY	781	GAACTCTTTGACGCTCCTCAGCCTGCCCTTGGAGCGCTGCAGCCTTAAGCCCAACAG	840
Db	1225	GAACTCTTTGACGCTCCTCAGCCTGCCCTTGGAGCGCTGCAGCCTTAAGCCCAACAG	1284
QY	841	CACAGTACCCAAAGGAGGAAAGCAACCTTAAGCGCGGAGAAAGTGAAGGCGCTTCCA	900
Db	1285	CACAGTACCCAAAGGAGGAAAGCAACCTTAAGCGCGGAGAAAGTGAAGGCGCTTCCA	1344
QY	901	ACGTTGATGCCCTTCTTCTTCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT	960
Db	1345	ACGTTGATGCCCTTCTTCTTCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT	1404
QY	961	AGGATGAGATTTGATTTATCCCTCCTCCCAACACCTAGGAAGTGAATCTTTTCTTTT	1020
Db	1405	AGGATGAGATTTGATTTATCCCTCCTCCCAACACCTAGGAAGTGAATCTTTTCTTTT	1464
QY	1021	TATTTTGTAGATGAGTCTTGTCTGTGCCCCAGCTGGAGTGCAGTGTGATCTCAG	1080
Db	1465	TATTTTGTAGATGAGTCTTGTCTGTGCCCCAGCTGGAGTGCAGTGTGATCTCAG	1524
QY	1081	CTTACTGCAACCTTGTCTCCCGGTTCAAGCATTTCTCCATCTCAGCCTCTCTGATAG	1140
Db	1525	CTTACTGCAACCTTGTCTCCCGGTTCAAGCATTTCTCCATCTCAGCCTCTCTGATAG	1584
QY	1141	CTGGAT 1147	
Db	1585	CTGGAT 1591	

RESULT 4  
BC050271  
LOCUS BC050271 4274 bp mRNA linear PRI 11-APR-2003  
DEFINITION Homo sapiens, similar to RIKEN cdna 5730551f12 gene, clone  
MGC:17391 IMAGE:3913714, mRNA, complete cds.  
ACCESSION BC050271  
VERSION BC050271.1 GI:29792138  
KEYWORDS MGC.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4274)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [md@paxil.stanford.edu](mailto:md@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 22 Row: 1 Column: 8  
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.

FEATURES  
source 1..4274  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:17391 IMAGE:3913714"  
/tissue\_type="Uterus, leiomyosarcoma"  
/clone\_lib="NIH\_MGC\_71"  
/lab\_host="DH10B"  
/notes="Vector: pCMV-SPORT6"  
93..2201  
/codon\_start=1  
/product="Similar to RIKEN cdna 5730551f12 gene"  
/protein\_id="AAH50271.1"  
/db\_xref="GI:29792139"  
/translation="MAQHPGERGLYCAHSHGASLRTLGSVDPEIPSPGSRDASG  
TAPNGTRCTEHSQKTPHNPALWLDSPGPPGPPDAEDDQSDSTSEEGSP  
VDDELAKENETGQKDNFSLTPSACNCOGTPIPEGVPYSEGNGSSNFCCHCTSP  
ALGEDLEBEYDDDESLKPSDFSRVSSSKKPPSRQRHRRPTKEDTREGRRDPRSP  
GHRLLKRSQADRRKGLGWLWABELCOLGQAGFWMLIELLVILVGEVETCGHLIYAC  
RQLKSDLDLFRVMGVTGRLGGWAQVMFQLSQGYFCYGVGLFTLRKLLALLLA  
LALFLGLQGLWRFVLGDLGRDKATWLFSLDPSALQRCGLTLRLDPRWRLVR  
IVQWGLELPWKQINROCNAPVAGRYCQPEEVARLITMAGVDELDNPFHLGV  
EATASDVLKAYRQLAVMHPDKNHPRAEAFKVLRAWDIVSNAEKREKEMKRM  
AENELSRVNEFLSLQDDDKENMTMCSRCQGHRRFEMDRPKSARTYACBENRHH  
PABEGDFWAESSLGLKITYFALMDKGYDITWAGCQRVIGSPDTHRVYHYSFSR  
IPTGRRGATTDAPADLQDFLSRIFQVPPGMPNGNFFAAPPQAPGAAAAKPNST  
VPKGAEPKRRKKVRPFQR"

CDS  
BASE COUNT 1065 a 1101 c 1143 g 965 t  
ORIGIN  
Query Match 100.0%; Score 1147; DB 9; Length 4274;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTGGTCAAGCAGAAATTAATAGGAGGGGAATGACCTGTAGCTAGTGGCGCTA 60  
Db 1295 GCCTTGGTCAAGCAGAAATTAATAGGAGGGGAATGACCTGTAGCTAGTGGCGCTA 1354  
QY 61 CTCCAGAGCTGAAGAGAGTGGCTGACTTGTACCATGGCTGGGCTCCTCAGGATGA 120  
Db 1355 CTCCAGAGCTGAAGAGAGTGGCTGACTTGTACCATGGCTGGGCTCCTCAGGATGA 1414



QY	301	GGAGTATGAGATGAACGAATGGCAGAGATGAGCTGAGCGGTGCTAGTAATGAGTTTCT	360
Db	745	GGAGTATGAGATGAACGAATGGCAGAGATGAGCTGAGCGGTGCTAGTAATGAGTTTCT	804
QY	361	GTCCAGCTCAGATGACCTCAGAGGCAATGAATACATATGATGTGAGCGGATGCCA	420
Db	805	GTCCAAGCTCAGATGACCTCAGAGGCAATGAATACATATGATGTGAGCGGATGCCA	864
QY	421	AGGAAGCATAGGAGGTTTCAATGGACCGGCACTTAAGAGTGCCAGATACCTGTGCTGA	480
Db	865	AGGAAGCATAGGAGGTTTCAATGGACCGGCACTTAAGAGTGCCAGATACCTGTGCTGA	924
QY	481	GTCTAATAGCTCATCTCTCAGGAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	540
Db	925	GTCTAATAGCTCATCTCTCAGGAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG	984
QY	541	CCTCAGATCACCCTACTTTGCATGATGATGTAAGAGGTGTATGACATCAGAGTGGC	600
Db	985	CCTCAGATCACCCTACTTTGCATGATGATGTAAGAGGTGTATGACATCAGAGTGGC	1044
QY	601	TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATCACATCTCAT	660
Db	1045	TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTATCACATCTCAT	1104
QY	661	TGGTTCCTCGGATCCAGGACACAGAGGCGGCGAGAGCCACCCAGATCCCTCTCTGC	720
Db	1105	TGGTTCCTCGGATCCAGGACACAGAGGCGGCGAGAGCCACCCAGATCCCTCTCTGC	1164
QY	721	TGATCTTCAGGATTTCTTGTAGTCGGATCTTTTCAAGTACCCCGAGGAGTGCCTCAT	780
Db	1165	TGATCTTCAGGATTTCTTGTAGTCGGATCTTTTCAAGTACCCCGAGGAGTGCCTCAT	1224
QY	781	GAACTCTTTGAGCTCCTCAGCTGCTGAGCGGCTGAGCGCTTGAAGCCCAACAG	840
Db	1225	GAACTCTTTGAGCTCCTCAGCTGCTGAGCGGCTGAGCGCTTGAAGCCCAACAG	1284
QY	841	CACAGTACCCAGGAGAACCAACCTTAAGCGCGGAGAAAGTGAAGGCGCTTCCA	900
Db	1285	CACAGTACCCAGGAGAACCAACCTTAAGCGCGGAGAAAGTGAAGGCGCTTCCA	1344
QY	901	ACGTTGATGCCCTTCTCTTCTCAATCAATGTGAGGAGTCAAAAGGCTGTAGCAC	960
Db	1345	ACGTTGATGCCCTTCTCTTCTCAATCAATGTGAGGAGTCAAAAGGCTGTAGCAC	1404
QY	961	AGGATGAGTTGATTTATCCCTCTCCCGCAACACTAGGAACCTGAATCTTTCTTT	1020
Db	1405	AGGATGAGTTGATTTATCCCTCTCCCGCAACACTAGGAACCTGAATCTTTCTTT	1464
QY	1021	TATTTTGTGAGTGGAGTCTTCTCTGTTCCCGAGCTGAGTGCAGTCTCAG	1080
Db	1465	TATTTTGTGAGTGGAGTCTTCTCTGTTCCCGAGCTGAGTGCAGTCTCAG	1524
QY	1081	CTTACTGCAACCTCTGCTCCCGGTTCAAGCAATCTCCCATCTCAGCTCTGAGTAG	1140
Db	1525	CTTACTGCAACCTCTGCTCCCGGTTCAAGCAATCTCCCATCTCAGCTCTGAGTAG	1584
QY	1141	CTGGGAT 1147	
Db	1585	CTGGGAT 1591	
RESULT 3			
AK055945			
LOCUS			
DEFINITION	AK055945	2343 bp mRNA linear PRI 01-AUG-2002	
		Homo sapiens cDNA FLJ31383 fis, clone NHMPC2000206, highly similar	
		to Bovine viral diarrhoea virus type 2 strain BVDV2-SD1630C	
		polyprotein gene.	
ACCESSION	AK055945		
VERSION	AK055945.1	GI:16550797	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	
AUTHORS	Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	
AUTHORS	Isogai, T., Otsuki, T. and Sugiyama, T.
TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES	Location/Qualifiers
source	1..2343
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="NHNP2000206"
	/cell_type="normal neural progenitor cells (NHNP958)"
	/clone_lib="NHNP2"
	/notes="cloning vector: pME18SFL3-primary culture, normal neural progenitor cells"
	113..1351
	/notes="unnamed protein product"
	/codon_start=1
	/protein_id="BAB71050.1"
	/db_xref="GI:16550798"
	/translation="MGVWTGRIGGHWQVFLSQFSGYCGVGLFTRFLKLLGALLLLA LALFLGLQGLWRFVLGDLRGWRDKATWFLSWDLPALRCUTLLRDSRPQWRLVR IVQWGLELPWYKQINRQGNAPVAGSGRYCQPEEVARLLTWAGVPEDELNPFVILGV EATASDELKAYRQVLAVMVHPDKNHPRAEAFKVLRAWDIVSNAERKPEYEMKRM ANELSRVNEFLSKLQDDLKEAMNTMCSRCQGHRRREMDEPESARYCAECNRLH PREGDPWAESSMLGLKITYPALMDGKYDITWAGCQRVGSPDTHRPVPHISFGSR IFTGRQRATPDAPADLQDLFLSRIFQVPPGQMPNGNFFAAPQAPAGAAAASKPNST VPKGEAKPKRRKKRRPFQR"
CDS	547 a 527 c 633 g 636 t
BASE COUNT	547 a 527 c 633 g 636 t
ORIGIN	
Query Match	100.0%; Score 1147; DB 9; Length 2343;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1147; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 GCCTTGGGTCAAGCAGAAATATTAATAGGCGAGGGAATGCACCTGTAGCTAGTGGCGCTA 60
Db	445 GCCTTGGGTCAAGCAGAAATATTAATAGGCGAGGGAATGCACCTGTAGCTAGTGGCGCTA 504
Qy	61 CTGCCAGCTCAAGCAGAAAGTGCCTGACTCTTGACCATGGCTGGGGTTCCTCAGGATGA 120
Db	505 CTGCCAGCTCAAGCAGAAAGTGCCTGACTCTTGACCATGGCTGGGGTTCCTCAGGATGA 564
Qy	121 GCTAAACCTTTCCATGTACTGGGGTTCAGGCCACAGCATCAGATGTTGAACTGAAGAA 180
Db	565 GCTAAACCTTTCCATGTACTGGGGTTCAGGCCACAGCATCAGATGTTGAACTGAAGAA 624
Qy	181 GGCCTATAGACAGCTGGCAGTGTGTTTCATCTCTGACAAAAATCATATCCCGGCTGA 240
Db	625 GGCCTATAGACAGCTGGCAGTGTGTTTCATCTCTGACAAAAATCATATCCCGGCTGA 684

Incyte Genomics, Inc. (US)	
FEATURES	Location/Qualifiers
source	1..1147
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/notes="Incyte ID No: 1459372CB1"
BASE COUNT	287 a 278 c 314 g 268 t
ORIGIN	
Query Match 100.0%; Score 1147; DB 6; Length 1147;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCCTTGGGTCAAGCAGAAATATTATAGCGAGGGAATGCACCTGTAGCTAGTGGGCGCTA 60
DB	1 GCCTTGGGTCAAGCAGAAATATTATAGCGAGGGAATGCACCTGTAGCTAGTGGGCGCTA 60
QY	61 CTGCCAGCCTGAAGAGAAAGTGGCTCGACTCTTGACCACTCTTGAGCCATGCTCGAGGATGA 120
DB	61 CTGCCAGCCTGAAGAGAAAGTGGCTCGACTCTTGACCACTCTTGAGCCATGCTCGAGGATGA 120
QY	121 GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACCTGAAGAA 180
DB	121 GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACCTGAAGAA 180
QY	181 GGCTATAGACAGTCGCAGTGTGTTTCATCTCTGACAAAATCATCATCCCCGGGCTGA 240
DB	181 GGCTATAGACAGTCGCAGTGTGTTTCATCTCTGACAAAATCATCATCCCCGGGCTGA 240
QY	241 GGAGGCCCTCAAGGTTTGGCAGCAGCTTGGGCAATTTGTCAGCAATGCTGAAAAGCGAAA 300
DB	241 GGAGGCCCTCAAGGTTTGGCAGCAGCTTGGGCAATTTGTCAGCAATGCTGAAAAGCGAAA 300
QY	301 GGAGTATGAGATGAACGAATGGCAGAGATGAGCTGAGCGGTGAGTAATGAGTTTCT 360
DB	301 GGAGTATGAGATGAACGAATGGCAGAGATGAGCTGAGCGGTGAGTAATGAGTTTCT 360
QY	361 GTCCAGCTCAAGATGACCTCAAGGAGCAATGAATACTATGATGTAGCCGATGCCA 420
DB	361 GTCCAGCTCAAGATGACCTCAAGGAGCAATGAATACTATGATGTAGCCGATGCCA 420
QY	421 AGGAAAGCATAGGAGGTTTCAATAGACCGGGAACCTTAAGAGTGCAGATACGTGCTGA 480
DB	421 AGGAAAGCATAGGAGGTTTCAATAGACCGGGAACCTTAAGAGTGCAGATACGTGCTGA 480
QY	481 GTCTAATAGCTGCATCTCTGCTGAGGAAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG 540
DB	481 GTCTAATAGCTGCATCTCTGCTGAGGAAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG 540
QY	541 CCTCAAGATCACCCTACTTTGCACTGATGATGTAAGAGGTTGATGACATCAGAGTGGG 600
DB	541 CCTCAAGATCACCCTACTTTGCACTGATGATGTAAGAGGTTGATGACATCAGAGTGGG 600
QY	601 TGGATCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCTATCAGATCTCATT 660
DB	601 TGGATCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCTATCAGATCTCATT 660
QY	661 TGGTTCTCGGATTTCCAGGCCACAGAGGCGGCGAGAGGCCACCCAGATGCCCTCCTGC 720
DB	661 TGGTTCTCGGATTTCCAGGCCACAGAGGCGGCGAGAGGCCACCCAGATGCCCTCCTGC 720
QY	721 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGG 780
DB	721 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCCAATGG 780
QY	781 GAACCTCTTTGAGCTCCTCAGCCTGCCCTGAGCGGCTGAGCCCTTAAGCCCAACAG 840
DB	781 GAACCTCTTTGAGCTCCTCAGCCTGCCCTGAGCGGCTGAGCCCTTAAGCCCAACAG 840
QY	841 CACAGTACCAAGGGAAGACCAACCTTAAGCGGCGGAAGAGTGAAGAGGCCCTTCCA 900
DB	841 CACAGTACCAAGGGAAGACCAACCTTAAGCGGCGGAAGAGTGAAGAGGCCCTTCCA 900
RESULT 2	
AX713973	
LOCUS	
Sequence 657 from Patent EP1293569.	
AX713973	
VERSION	
AX713973.1 GI:29888901	
KEYWORDS	
SOURCE	
ORGANISM	
Homo sapiens (human)	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	
AUTHORS	
1	
Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,	
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,	
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and	
Masuno,Y.	
Full-length cDNAs	
Patent: EP 1293569-A 657 19-MAR-2003;	
Helix Research Institute (JP); Research Association for	
Biotechnology (JP)	
FEATURES	
source	
1..2343	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
BASE COUNT	547 a 527 c 633 g 636 t
ORIGIN	
Query Match 100.0%; Score 1147; DB 6; Length 2343;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCCTTGGGTCAAGCAGAAATATTATAGCGAGGGAATGCACCTGTAGCTAGTGGGCGCTA 60
DB	445 GCCTTGGGTCAAGCAGAAATATTATAGCGAGGGAATGCACCTGTAGCTAGTGGGCGCTA 504
QY	61 CTGCCAGCCTGAAGAGAAAGTGGCTCGACTCTTGACCACTGCTGGGTTCTCTGAGGATGA 120
DB	505 CTGCCAGCCTGAAGAGAAAGTGGCTCGACTCTTGACCACTGCTGGGTTCTCTGAGGATGA 564
QY	121 GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACCTGAAGAA 180
DB	565 GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAACCTGAAGAA 624
QY	181 GGCCTATAGACAGCTGGCAGTGTGTTTCATCTCTGACAAAATCATCATCCCCGGGCTGA 240
DB	625 GGCCTATAGACAGCTGGCAGTGTGTTTCATCTCTGACAAAATCATCATCCCCGGGCTGA 684
QY	241 GGAGCCCTCAAGGTTTTCGAGCAGCTTGGGACATTTGTGACCAATGCTGAAAAGCGAAA 300
DB	685 GGAGCCCTCAAGGTTTTCGAGCAGCTTGGGACATTTGTGACCAATGCTGAAAAGCGAAA 744

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 1, 2003, 07:30:31 ; Search time 4369 Seconds  
(without alignments)  
10740.065 Million cell updates/sec  
Title: US-10-049-742-22  
Perfect score: 1147  
Sequence: 1 gccttggtgaagcagaata.....gcctcctgagtagctgggat 1147

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pin.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1147	100.0	1147	6	AX081438	AX081438 Sequence
2	1147	100.0	2343	6	AX713973	AX713973 Sequence
3	1147	100.0	2343	9	AK055945	AK055945 Homo sapi
4	1147	100.0	4274	9	BC050271	BC050271 Homo sapi
5	963.6	84.0	190201	2	AC129837	AC129837 Papio anu
6	883.4	77.0	1818	9	BC016941	BC016941 Homo sapi
7	867.2	75.6	953	9	AF351784	AF351784 Homo sapi
8	834	72.7	2402	4	AY027882	AY027882 Bos tauru
9	834	72.7	3159	4	AY027881	AY027881 Bos tauru
10	824	71.8	2692	4	AF308815	AF308815 Bos tauru
11	793.6	69.2	2402	10	BC011146	BC011146 Mus muscu
12	781.8	68.2	2392	10	AF351783	AF351783 Rattus no
13	765.8	66.8	263371	2	AC098454	AC098454 Rattus no
14	743	64.8	754	9	AF141342	AF141342 Homo sapi
15	661.2	57.6	247309	2	AC097837	AC097837 Rattus no
16	628.2	54.8	747	6	BD020639	BD020639 Novel gen
17	628.2	54.8	747	6	BD100577	BD100577 Novel gen
18	453	39.5	118052	9	AC023055	AC023055 Homo sapi
19	400.8	34.9	1253	14	AF268178	AF268178 Bovine vi
20	399	34.8	1258	14	AF268176	AF268176 Bovine vi
21	360.2	31.4	255930	2	AC123227	AC123227 Rattus no
22	354.2	30.9	198493	2	AC112674	AC112674 Mus muscu
23	354.2	30.9	250277	2	AC122757	AC122757 Mus muscu
24	352.8	30.8	1004	14	AF268171	AF268171 Bovine vi
25	352.8	30.8	1233	14	AF268172	AF268172 Bovine vi
26	337.2	29.4	927	14	BDU43603	BDU43603 Border dise
27	303.6	26.5	873	14	BDU43602	BDU43602 Border dise
28	298.6	26.0	194221	10	AC113998	AC113998 Mus muscu
29	297	25.9	717	14	BVD125IN1	BVD125IN1 Bovine vira
30	294.6	25.7	1229	14	AF268180	AF268180 Bovine vi
31	294.6	25.7	13196	14	AB078951	AB078951 Bovine vi
32	294.6	25.7	13203	14	AB078952	AB078952 Bovine vi
33	285.2	24.9	752	14	AF268177	AF268177 Bovine vi
34	279.6	24.4	711	14	BVD125IN2	BVD125IN2 Bovine vira
35	278.6	24.3	852	14	AF268179	AF268179 Bovine vi
36	272.8	23.8	12602	14	AF144617	AF144617 Pestiviru
37	251.4	21.9	12119	6	AX057315	AX057315 Sequence
38	251.4	21.9	12573	14	BVDCG	BVDCG
39	251.4	21.9	12578	14	BV1133738	BV1133738 Bovine vi
40	251.4	21.9	12734	6	AR179057	AR179057 Sequence
41	251.4	21.9	12734	14	AF268278	AF268278 Pestiviru
42	251.4	21.9	14078	6	AR123006	AR123006 Sequence
43	251.4	21.9	14078	6	AR214645	AR214645 Sequence
44	251.4	21.9	14578	6	AR094147	AR094147 Sequence
45	251.2	21.9	1114	14	AF268175	AF268175 Bovine vi

ALIGNMENTS

RESULT 1  
AX081438  
LOCUS AX081438  
DEFINITION Sequence 22 from Patent WO0109178.  
ACCESSION AX081438  
VERSION AX081438.1 GI:13170261  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Yue,H., Bandman,O., Tang,Y.T., Baughn,M.R., Azimzai,Y. and Lu,D.A.  
TITLE Human chaperone proteins  
JOURNAL Patent: WO 0109178-A 22 08-FEB-2001;

Query Match 21.9%; Score 251.4; DB 21; Length 12842;  
Best Local Similarity 94.2%; Pred. No. 3.3e-66;  
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 404 ATGTGTAGCCGATGCCAGGAAGCATAGGAGGTTTGAATGGACCGGAACCTTAAGAGT 463  
DB 4962 ATGTGCACCCGATGCCAGGAAGCATAGGAGGTTTGAATGGACCGGAACCTTAAGAGT 5021

QY 464 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTGCTGAGGAAGGAGACTTTTGGGCA 523  
DB 5022 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTGCTGAGGAAGGAGTACTTTTGGGCA 5081

QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTTGCACTGATGGAAAGGTGTAT 583  
DB 5082 GAGTCGACCATGTTGGGCTCAAAATCACCTACTTTTGGCTGATGGATGAAAGGTGTAT 5141

QY 584 GACATCACAGAGTGGGCTGGATGCCAGCGTGTAGTATCTCCAGATACCCACAGAGTC 643  
DB 5142 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCAGATACCCACAGAGTC 5201

QY 644 CCTATCACATCTCATTTGGTTCTCGGATTCAGGCA 680  
DB 5202 CCTGTGCATCTCATTTGGTTCTACGGATGCCTTTCA 5238

RESULT 15  
AAZ36211  
ID AAZ36211 standard; DNA; 13198 BP.  
XX  
AC AAZ36211;  
XX  
DT 11-FEB-2000 (first entry)  
XX  
DE Nucleotide sequence of functional HCV-BVDV chimera from pCBV/p7/IRIS-pac.  
XX  
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;  
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;  
KW bovine viral diarrhea virus; NADL; vaccine; ss.  
XX  
OS Chimeric - Hepatitis C virus.  
OS Chimeric - Bovine viral diarrhea virus.  
XX  
PN WO9953366-A1.  
XX  
PD 04-NOV-1999.  
XX  
PF 23-APR-1999; 99WO-US08850.  
XX  
PR 24-APR-1998; 98US-0082964.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Rice CW, Frolov I, McBride MS;  
XX  
DR WPI; 2000-013359/01.  
XX  
PT Chimeric viral RNA, used in vaccine against BVDV -  
XX  
PS Example 3; Fig 24; 108pp; English.  
XX

The present sequence represents a functional Hepatitis C virus (HCV)-bovine viral diarrhea virus (BVDV) chimeric virus of the invention, expressing a dominant selectable marker conferring resistance to puromycin. The specification describes chimeric viral RNA comprising a 5' nontranslated region (5'NTR); an open reading frame (ORF) region; and a 3' NTR; where at least one of the regions is chimeric and comprises a nucleotide sequence from a pestivirus in operable linkage with a heterologous nucleotide sequence, preferably from HCV. The chimeric viral RNA is replication-competent. The chimeric viral RNA can be used in a method for identifying compounds having antiviral activity against HCV. When the pestivirus viral nucleotide sequence is from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be

CC used in a vaccine against BVDV.

XX  
SQ Sequence 13198 BP; 3735 A; 3097 C; 3533 G; 2833 T; 0 other;  
Query Match 21.9%; Score 251.4; DB 21; Length 13198;  
Best Local Similarity 94.2%; Pred. No. 3.3e-66;  
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 404 ATGTGTAGCCGATGCCAGGAAGCATAGGAGGTTTGAATGGACCGGAACCTTAAGAGT 463  
DB 4089 ATGTGCACCCGATGCCAGGAAGCATAGGAGGTTTGAATGGACCGGAACCTTAAGAGT 4148

QY 464 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTGCTGAGGAAGGAGACTTTTGGGCA 523  
DB 4149 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTGCTGAGGAAGGAGTACTTTTGGGCA 4208

QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTTGCACTGATGGAAAGGTGTAT 583  
DB 4209 GAGTCGACCATGTTGGGCTCAAAATCACCTACTTTTGGCTGATGGATGAAAGGTGTAT 4268

QY 584 GACATCACAGAGTGGGCTGGATGCCAGCGTGTAGTATCTCCAGATACCCACAGAGTC 643  
DB 4269 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCAGATACCCACAGAGTC 4328

QY 644 CCTATCACATCTCATTTGGTTCTCGGATTCAGGCA 680  
DB 4329 CCTGTGCATCTCATTTGGTTCTACGGATGCCTTTCA 4365

Search completed: December 1, 2003, 07:37:13  
Job time : 384 secs

QY 404 ATGTGTAGCCGATGCCAAGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT 463  
DB 4994 ATGTGCAGCCGATGCCAGGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT 5053  
QY 464 GCCGATACCTGTGCTGAGTGTAAATAGGCTGCATCTCTGTGAGGAAGGAGACTTTTGGGCA 523  
DB 5054 GCCGATACCTGTGCTGAGTGTAAATAGGCTGCATCTCTGTGAGGAAGGAGTGACTTTTGGGCA 5113  
QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTGCACTGTGAGTGGAAAGGTGTAT 583  
DB 5114 GAGTCGAGCATGTTGGGCTCAAAATCACCTACTTTGCGCTGATGGATGGAAAGGTGTAT 5173  
QY 584 GACATCACAGAGTGGGCTGGATGCCAGCGTGTAGTATCTCTCCAGATACCCACAGAGTC 643  
DB 5174 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCAGATACCCACAGAGTC 5233  
QY 644 CCCTATCACATCTCAATTTGGTTCTCGGATTCAGGCA 680  
DB 5234 CCTTGTACATCTCAATTTGGTTTACGGATGCTTTCA 5270

RESULT 13  
ABA95615  
ID ABA95615 standard; DNA; 12734 BP.  
XX AC ABA95615;  
XX AC ABA95615;  
DT 21-MAR-2002 (first entry)  
XX Chimeric BVDV/HCV NS3-wt sequence.  
DE Pestivirus; Npro; protease; NS3; screening; ds.  
KW Chimeric - Bovine viral diarrhea virus.  
OS Chimeric - Hepatitis C virus.  
XX US6326137-B1.  
XX 04-DEC-2001.  
XX 25-JUN-1999; 99US-0344456.  
XX 25-JUN-1999; 99US-0344456.  
XX (SCHE ) SCHERING CORP.  
XX Hong Z, Lai VCH, Lau JYN;  
XX WPI; 2002-121103/16.  
XX Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)  
PT pestivirus genome where the Npro protease gene is replaced with NS3  
PT protease gene, useful for in vivo screening of compounds which inhibit  
PT HCV infection -  
XX  
XX Example 2; Columns 17-28; 20pp; English.

XX The present invention relates to a nucleic acid construct encoding a  
CC chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct  
CC comprises a pestivirus genome where a Npro pestivirus protease gene is  
CC replaced with a gene encoding a functional HCV NS3 protease. Furthermore,  
CC each junction site recognised by the Npro protease is replaced with a  
CC junction site recognised by the HCV NS3 protease. The construct is useful  
CC for screening compounds that inhibit HCV in vivo by inhibiting HCV  
CC protease, where screening may be in cell culture or in an animal model.  
CC The present sequence is a chimeric clone of BVDV (bovine viral diarrhea  
CC virus)/HCV NS3-wt, which was used to illustrate the present invention.  
XX  
XX Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;  
Query Match 21.9%; Score 251.4; DB 24; Length 12734;  
Best Local Similarity 94.2%; Pred. No. 3.3e-66;  
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 404 ATGTGTAGCCGATGCCAAGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT 463  
DB 5150 ATGTGCAGCCGATGCCAGGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT 5209  
QY 464 GCCGATACCTGTGCTGAGTGTAAATAGGCTGCATCTCTGTGAGGAAGGAGACTTTTGGGCA 523  
DB 5210 GCCGATACCTGTGCTGAGTGTAAATAGGCTGCATCTCTGTGAGGAAGGAGTGACTTTTGGGCA 5269  
QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACCTACTTTGCACTGTGAGTGGAAAGGTGTAT 583  
DB 5270 GAGTCGAGCATGTTGGGCTCAAAATCACCTACTTTGCGCTGATGGATGGAAAGGTGTAT 5329  
QY 584 GACATCACAGAGTGGGCTGGATGCCAGCGTGTAGTATCTCTCCAGATACCCACAGAGTC 643  
DB 5330 GATATCACAGAGTGGGCTGGATGCCAGCGTGTGGGAATCTCCAGATACCCACAGAGTC 5389  
QY 644 CCCTATCACATCTCAATTTGGTTCTCGGATTCAGGCA 680  
DB 5390 CCTTGTACATCTCAATTTGGTTTACGGATGCTTTCA 5426

RESULT 14  
AAZ36203  
ID AAZ36203 standard; DNA; 12842 BP.  
XX AC AAZ36203;  
XX AC AAZ36203;  
DT 11-FEB-2000 (first entry)  
XX Nucleotide sequence of the prototype HCV-BVDV chimera.  
KW Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;  
KW 5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;  
KW bovine viral diarrhea virus; NADL; vaccine; ss.  
XX Chimeric - Hepatitis C virus.  
OS Chimeric - Bovine viral diarrhea virus.  
XX WO9955366-A1.  
XX 04-NOV-1999.  
XX 23-APR-1999; 99WO-US08850.  
XX 24-APR-1998; 98US-0082964.  
XX (UNIW ) UNIV WASHINGTON.  
XX Rice CM, Frolov I, McBride MS;  
XX WPI; 2000-013359/01.  
XX Chimeric viral RNA, used in vaccine against BVDV -  
XX Example 5; Fig 19; 108pp; English.  
XX The present sequence represents the prototype Hepatitis C virus  
CC (HCV)-bovine viral diarrhea virus (BVDV) chimeric virus of the  
CC invention. The sequence contains the adapted HCV 5'NTR from  
CC 5'NTR/R.3orig and tandem 3'NTR elements from HCV followed by BVDV.  
CC The specification describes chimeric viral RNA comprising a  
CC 5' nontranslated region (5'NTR); an open reading frame (ORF) region;  
CC and a 3' NTR; where at least one of the regions is chimeric and  
CC comprises a nucleotide sequence from a pestivirus in operable linkage  
CC with a heterologous nucleotide sequence, preferably from HCV. The  
CC chimeric viral RNA is replication-competent. The chimeric viral RNA  
CC can be used in a method for identifying compounds having antiviral  
CC activity against HCV. When the pestivirus viral nucleotide sequence is  
CC from bovine viral diarrhea virus (BVDV), the chimeric viral RNA can be  
CC used in a vaccine against BVDV.  
XX Sequence 12842 BP; 4034 A; 2612 C; 3282 G; 2914 T; 0 other;

QY	524	GAGTCAAGCATGTTGGGCCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTAT	583
DB	4655	GAGTCGAGCATGTTGGGCCTCAAAATCACCTACTTTGCGCTGATGGATGGAAGGTGTAT	4714
QY	584	GACATCACAGAGTGGGCTGGATGCCACGCTGTAGGTATCTCCCCAGATACCCACAGATC	643
DB	4715	GATATCACAGAGTGGGCTGGATGCCACGCTGTGGGAATCTCCCCAGATACCCACAGATC	4774
QY	644	CCCTATCACATCTCATTTTGGTTCTCGGATTCAGGCA	680
DB	4775	CTTGTGCATCTCATTTTGGTTCTCGGATTCAGGCA	4811
RESULT 12			
AAZ36196			
ID	AAZ36196	standard; cDNA; 12578 BP.	
XX			
AC	AAZ36196;		
XX			
DT	11-FEB-2000	(first entry)	
XX			
DE	Nucleotide	sequence of infectious BVDV NADL protein.	
XX			
KW	Pseudorevertant; RNA virus; chimera; BVDV; HCV; replication-competent;		
KW	5' nontranslated region; 5'NTR; 3' NTR; pestivirus; antiviral;		
KW	bovine viral diarrhoea virus; NADL; vaccine; ss.		
XX			
OS	Bovine viral diarrhoea virus.		
XX			
EH	Key	Location/Qualifiers	
FT	CDS	386..12352	
FT		/*tag= a	
FT		/product= "NADL protein"	
XX			
PN	WO9955366-A1.		
XX			
PD	04-NOV-1999.		
XX			
PF	23-APR-1999;	99WO-US08850.	
XX			
PR	24-APR-1998;	98US-0082964.	
XX			
PA	(UNIW )	UNIV WASHINGTON.	
XX			
PI	Rice CM, Frollov I, McBride MS;		
XX			
DR	WPI: 2000-013359/01.		
DR	P-PSDB; AAY53616.		
XX			
PT	Chimeric viral RNA, used in vaccine against BVDV .		
XX			
PS	Disclosure; Fig 11; 108pp; English.		
XX			
CC	The present sequence encodes the NADL protein of bovine viral diarrhoea		
CC	virus (BVDV). The sequence is used in the course of the invention, to		
CC	produce chimeric RNA viruses. The specification describes chimeric viral		
CC	RNA comprising a 5' nontranslated region (5'NTR); an open reading frame		
CC	(ORF) region; and a 3' NTR, where at least one of the regions is chimeric		
CC	and comprises a nucleotide sequence from a pestivirus in operable linkage		
CC	with a heterologous nucleotide sequence, preferably from HCV. The		
CC	chimeric viral RNA is replication-competent. The chimeric viral RNA		
CC	can be used in a method for identifying compounds having antiviral		
CC	activity against HCV. When the pestivirus viral nucleotide sequence is		
CC	from bovine viral diarrhoea virus (BVDV), the chimeric viral RNA can be		
CC	used in a vaccine against BVDV.		
XX			
SQ	Sequence 12578 BP; 4049 A; 2527 C; 3233 G; 2769 T; 0 other;		
Query Match	21.9%;	Score 251.4; DB 21; Length 12578;	
Best Local Similarity	94.2%;	Pred. No. 3.2e-66;	
Matches 261; Conservative	0; Mismatches 16; Indels 0; Gaps 0;		

QY	404	ATGTGTAGCCGATGCCAAGCAAGCATAGGAGTTTGAATGACCGGAACTTAGAGT	463
DB	4535	ATGTGACCGCATGCCAGGAAGCATAGAGTTTGAATGACCGGAACTTAGAGT	4594
QY	464	GCCAGATAGTGTGCTGAGTGTAATAGGCTGCATCTCTCAGGAAGGACATTTTGGGCA	523
DB	4595	GCCAGATAGTGTGCTGAGTGTAATAGGCTGCATCTCTCAGGAAGGACATTTTGGGCA	4654

```
XX SQ Sequence 7453 BP; 1980 A; 1633 C; 1692 G; 2147 T; 1 other;
Query Match 39.5%; Score 453; DB 22; Length 7453;
Best Local Similarity 100.0%; Pred. No. 1.5e-128;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 695 AGAGCCACCCAGATGCCCTCCTGCTGATCTTCAGGATTTCTTGAGTCGGATCTTTCAA 754
Db 3315 AGAGCCACCCAGATGCCCTCCTGCTGATCTTCAGGATTTCTTGAGTCGGATCTTTCAA 3374

QY 755 GTACCCCCAGGCGAGATGCCCAATGGGAATCTTCTTCAGCTCCTCAGCCCTGCCCTTGA 814
Db 3375 GTACCCCCAGGCGAGATGCCCAATGGGAATCTTCTTCAGCTCCTCAGCCCTGCCCTTGA 3434

QY 815 GCCCTCAGCCTCTAAGCCCAACAGACACAGTACCCAAAGGAGAGCAAAACCTTAAGCGG 874
Db 3435 GCCCTCAGCCTCTAAGCCCAACAGACACAGTACCCAAAGGAGAGCAAAACCTTAAGCGG 3494

QY 875 CGGAAGAAAGTGAGGAGGCCCTTCCAACTGTGATGCCCTTCTCTTCTCCTCAAAATCAATG 934
Db 3495 CGGAAGAAAGTGAGGAGGCCCTTCCAACTGTGATGCCCTTCTCTTCTCCTCAAAATCAATG 3554

QY 935 TCAGGAGTCAAAAGGCTGTAGCACAGGATGGAGTTGATTTATCCTCTCCTCCCAAC 994
Db 3555 TCAGGAGTCAAAAGGCTGTAGCACAGGATGGAGTTGATTTATCCTCTCCTCCCAAC 3614

QY 995 ACCTPAGAACTGAATCTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1054
Db 3615 ACCTPAGAACTGAATCTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3674

QY 1055 GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTCTGTCTCCCGGTTCAAGCAA 1114
Db 3675 GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTCTGTCTCCCGGTTCAAGCAA 3734

QY 1115 TTCTCCATCTCAGCTCCTGAGTGGAT 1147
Db 3735 TTCTCCATCTCAGCTCCTGAGTGGAT 3767

RESULT 10
ABL55114
ID ABL55114 standard; DNA; 884 BP.
XX AC ABL55114;
XX DT 08-OCT-2002 (first entry)
XX DE Human cDNA sequence #2 from clone HNTPB82.
XX KW Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
KW vasotrophic; cerebroprotective; neurotropic; neuroprotective; antibacterial;
KW virucide; fungicide; ophthalmological; vulnerary; gene therapy; ELISA;
KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; food additive; ss.
XX OS Homo sapiens.
XX PH Location/Qualifiers
FT CDS 272..883
FT FT /*tag= a
FT FT /note= "No stop codon given"
FT sig_peptide 272..400
FT FT /*tag= b
FT mat_peptide 401..883
FT FT /*tag= c
XX WO200222638-A1.
XX PN 21-MAR-2002.
XX PD
```

```
XX PF 17-JAN-2001; 2001WO-US01386.
XX PR 12-SEP-2000; 2000US-232104P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fisceila M;
XX Ni J;
XX WPI; 2002-258041/30.
XX DR P-PSDB; ABB77045.
XX PT New nucleic acid molecules encoding 22 human secreted proteins for
XX PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative
XX PT disorders, and cardiovascular disorders, and used as food additives or
XX PT preservatives -
XX PS Disclosure; Page 466-467; 526pp; English.
XX PS The sequence represents a cDNA sequence of the invention, isolated
XX CC from human clone ID HNPB82. The invention relates to novel
XX CC isolated nucleic acid molecules encoding 22 human secreted proteins. The
XX CC proteins of the invention have immunosuppressive, antirheumatic,
XX CC antitumetic, antiproliferative, cytostatic, cardiant, vasotrophic,
XX CC cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide,
XX CC fungicide, ophthalmological, and vulnerary activity. The polynucleotides
XX CC may have a use in gene therapy. The polynucleotides and polypeptides
XX CC encoded by them are used to prevent, treat or ameliorate a medical
XX CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
XX CC chickens or sheep. The polynucleotides and polypeptides are also used in
XX CC diagnosing a pathological condition or susceptibility to a pathological
XX CC condition. The antibodies to the proteins can also be used in alleviating
XX CC symptoms associated with the disorders and in diagnostic immunoassays
XX CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
XX CC Disorders which are diagnosed or treated include autoimmune diseases,
XX CC hyperproliferative disorders, cardiovascular disorders, cerebrovascular
XX CC disorders, angiogenesis, nervous system disorders, infections caused by
XX CC bacteria, viruses and fungi and ocular disorders. The polypeptides can
XX CC also be used to aid wound healing and epithelial cell proliferation. The
XX CC polypeptides can also be used as a food additive or preservative.
XX SQ Sequence 884 BP; 187 A; 178 C; 291 G; 228 T; 0 other;

Query Match 23.8%; Score 272.8; DB 24; Length 884;
Best Local Similarity 95.6%; Pred. No. 1.9e-73;
Matches 302; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 1 GCCTTGGGTCAACGACAGAAATATTATAGGCGAGGGGAATGCACCTGTAGCTAGTGGCGCTA 60
Db 562 GCCTTGGGTCAACGACAGAAATATTATAGGCGAGGGGAATGCACCTGTAGCTAGTGGCGCTA 621

QY 61 CTGCCAGCCTGAGAGAGAGTGGCTCGACTTTGACCTTTCACCTCGCTGGGTTCTCTGAGGATGA 120
Db 622 CTGCCAGCCTGAGAGAGAGTGGCTCGACTTTGACCTTTCACCTCGCTGGGTTCTCTGAGGATGA 681

QY 121 GCTAAACCCCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATCTTCAACTGAAGAA 180
Db 682 GCTAAACCCCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATCTTCAACTGAAGAA 741

QY 181 GGCTTATAGACAGCTGGCAGTGTGTTTCTATCTGTACAAAATCATCATCCCCGGGCTGA 240
Db 742 GGCTTATAGACAGCTGGCAGTGTGTTTCTATCTGTACAAAATCATCATCCCCGGGCTGA 801

QY 241 GGAGGCTTCAAGG-TTTTCCGAGCAGCTTT-GGACATTTCTAGCAATGTCTGAAAAGCGA 298
Db 802 GGAGGCTTCAAGGTTTTCGAGCAGCTTTGGGACATTTGTGAGCAATGTCTGAAAAGCG 861

QY 299 AGGAGTATCAGATGA 314
Db 862 AAAGGAGTATGAGGA 877
```



PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 05-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234597.  
PR 26-SEP-2000; 2000US-0234598.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX WPI; 2001-465570/50.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen  
PT is used in preventing, treating or ameliorating a medical condition -  
XX  
XX Disclosure; SEQ ID NO 8949; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.

PR	08-NOV-2000;	2000US-0246478.	CC	and neurological disorders (Alzheimer's disease and Parkinson's disease),
PR	08-NOV-2000;	2000US-0246523.	CC	endocrine disorders (Addison's disease), gastrointestinal disorders
PR	08-NOV-2000;	2000US-0246524.	CC	(inflammatory disorders), liver disorders (biliary liver cirrhosis),
PR	08-NOV-2000;	2000US-0246525.	CC	pancreatic and gall bladder disorders, disorders of the large intestine,
PR	08-NOV-2000;	2000US-0246526.	CC	developmental and inherited disorders, diseases at the cellular level,
PR	08-NOV-2000;	2000US-0246527.	CC	and wound healing and epithelial cell proliferation. (I) or (II) is
PR	08-NOV-2000;	2000US-0246528.	CC	useful to prevent skin aging, for preventing hair loss, to maintain
PR	08-NOV-2000;	2000US-0246532.		
PR	08-NOV-2000;	2000US-0246609.		
PR	08-NOV-2000;	2000US-0246610.		
PR	08-NOV-2000;	2000US-0246611.		
PR	08-NOV-2000;	2000US-0246613.		
PR	17-NOV-2000;	2000US-0249207.		
PR	17-NOV-2000;	2000US-0249208.		
PR	17-NOV-2000;	2000US-0249209.		
PR	17-NOV-2000;	2000US-0249210.		
PR	17-NOV-2000;	2000US-0249211.		
PR	17-NOV-2000;	2000US-0249212.		
PR	17-NOV-2000;	2000US-0249213.		
PR	17-NOV-2000;	2000US-0249214.		
PR	17-NOV-2000;	2000US-0249215.		
PR	17-NOV-2000;	2000US-0249216.		
PR	17-NOV-2000;	2000US-0249217.		
PR	17-NOV-2000;	2000US-0249218.		
PR	17-NOV-2000;	2000US-0249244.		
PR	17-NOV-2000;	2000US-0249245.		
PR	17-NOV-2000;	2000US-0249264.		
PR	17-NOV-2000;	2000US-0249265.		
PR	17-NOV-2000;	2000US-0249297.		
PR	17-NOV-2000;	2000US-0249299.		
PR	17-NOV-2000;	2000US-0249300.		
PR	01-DEC-2000;	2000US-0250160.		
PR	01-DEC-2000;	2000US-0250391.		
PR	05-DEC-2000;	2000US-0251030.		
PR	05-DEC-2000;	2000US-0251988.		
PR	05-DEC-2000;	2000US-0256719.		
PR	06-DEC-2000;	2000US-0251479.		
PR	08-DEC-2000;	2000US-0251856.		
PR	08-DEC-2000;	2000US-0251868.		
PR	08-DEC-2000;	2000US-0251869.		
PR	08-DEC-2000;	2000US-0251989.		
PR	08-DEC-2000;	2000US-0251990.		
PR	11-DEC-2000;	2000US-0254097.		
PR	05-JAN-2001;	2001US-0259678.		
XX				
PA	(HUMA-) HUMAN GENOME SCI INC.			
XX				
PI	Rosen CA, Barash SC, Ruben SM;			
XX				
DR	WPI; 2001-476223/51.			
XX				
PT	Novel isolated prostate gland related polypeptide useful for diagnosis			
PT	and treatment of disorders of prostate such as prostatodystonia,			
PT	prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia			
PT				
XX				
PS	Claim 1; SEQ ID No 339; 512pp; English.			
XX				
CC	The invention relates to novel isolated prostate gland related nucleic			
CC	acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,			
CC	prognosis, prevention, and/or treatment of diseases and/or disorders of			
CC	the prostate such as acute non-bacterial prostatitis, chronic non-			
CC	bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,			
CC	prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic			
CC	hypertrophy or hyperplasia, and prostate neoplastic disorders, including			
CC	adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and			
CC	squamous cell carcinomas. (I), (II) and antibody to (I) are useful for			
CC	diagnosing and treating reproductive system disorders (Paget's disease),			
CC	autoimmune disorders (systemic lupus erythematosus, rheumatoid			
CC	arthritis), blood-related disorders (sickle cell anaemia),			
CC	hyperproliferative disorders, urinary system disorders			
CC	(glomerulonephritis), cardiovascular disorders (arrhythmias),			
CC	respiratory disorders, musculoskeletal system disorders, neural activity			

CC	and neurological disorders (Alzheimer's disease and Parkinson's disease),
CC	endocrine disorders (Addison's disease), gastrointestinal disorders
CC	(inflammatory disorders), liver disorders (biliary liver cirrhosis),
CC	pancreatic and gall bladder disorders, disorders of the large intestine,
CC	developmental and inherited disorders, diseases at the cellular level,
CC	and wound healing and epithelial cell proliferation. (I) or (II) is
CC	useful to prevent skin aging, for preventing hair loss, to maintain
	Query Match 39.5%; Score 453; DB 22; Length 7453;
	Best Local Similarity 100.0%; Pred. No. 1.5e-128;
	Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	695 AGAGCCACCCAGATGCCCTCTCTGCTGATCTTTCAGGATTTCTTTCAGTGGATCTTCAA 754
Db	3315 AGAGCCACCCAGATGCCCTCTCTGCTGATCTTTCAGGATTTCTTTCAGTGGATCTTCAA 3374
Qy	755 GTACCCCGAGGGCAGATGCCCAATGGAACTCTTTTGCAGCTCTCTCAGCTGCCCTTGA 814
Db	3375 GTACCCCGAGGGCAGATGCCCAATGGAACTCTTTTGCAGCTCTCTCAGCTGCCCTTGA 3434
Qy	815 GCGCTGTGAGCTTAAAGCCCAACAGCACAGTACCCAGGGAGAGGCAACCTAAGCGG 874
Db	3435 GCGCTGTGAGCTTAAAGCCCAACAGCACAGTACCCAGGGAGAGGCAACCTAAGCGG 3494
Qy	875 CGGAAGAAAGTGAAGAGGCGCTTCCAAAGTGTGATGCCCTTCTCTTCTCAATCAATG 934
Db	3495 CGGAAGAAAGTGAAGAGGCGCTTCCAAAGTGTGATGCCCTTCTCTTCTCAATCAATG 3554
Qy	935 TCAGGGAGTCAAAAGGGCTGTAGCAGAGTGGAGTTGATTTATCCCTCTCTCCCCAAC 994
Db	3555 TCAGGGAGTCAAAAGGGCTGTAGCAGAGTGGAGTTGATTTATCCCTCTCTCCCCAAC 3614
Qy	995 ACCTAGAACTGAATCTTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1054
Db	3615 ACCTAGAACTGAATCTTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3674
Qy	1055 GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTCTGTCTCCGGGTTCAAGCA 1114
Db	3675 GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTCTGTCTCCGGGTTCAAGCA 3734
Qy	1115 TTCTCCCATCTCAGCTCTCTGAGTAGTGGAT 1147
Db	3735 TTCTCCCATCTCAGCTCTCTGAGTAGTGGAT 3767
RESULT 9	
AAU06261	
ID	AAU06261 standard; DNA; 7453 BP.
XX	
AC	AAU06261;
XX	
DT	21-NOV-2001 (first entry)
XX	
DE	Human reproductive system related antigen DNA SEQ ID NO: 8949.
XX	
KW	Human; reproductive system related antigen; reproductive system disorder;
KW	cancer; gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200155320-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US01339.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	24-FEB-2000; 2000US-0184664.
PR	02-MAR-2000; 2000US-0186350.
PR	16-MAR-2000; 2000US-0189874.
PR	17-MAR-2000; 2000US-0190076.
PR	18-APR-2000; 2000US-0198123.

Db	397	CTGAGGAAGGAGACTTTTGGGCAGAGTCAAGCATGTTGGGCCTCAAGATCACCTACTTTG	456
Qy	561	CACATGATGGATGGAAGGTGTATGACATCAGAGTGGCTGGATGCCAG-CGTGTAGGT	619
Db	457	CACATGATGGATGGAAGGTGTATGACATCAGAGTGGCTGGATGCCAGCCGTGTAGGT	516
Qy	620	ATCTCCCGAGATACCCACAGAGTCCCTATCACATCTCATTTGGTTCTCGGATTCAGGC	679
Db	517	ATCTCCCGAGATACCCACAGAGTCCCTATCACATCTCATTTGGTTCTCGGATTCAGGC	576
Qy	680	ACCAGAGGGCGCAGAGAGCCACCCAGATGCCCTCTGCTGATCTTCAGGATTTCTTG	739
Db	577	ACCAGAGGGCGCAGAGAGCCACCCAGATGCCCTCTGCTGATCTTCAGGATTTCTTG	636
Qy	740	AGTCGGATCTTTCAAGTACCCCGAGGACATGCCCAATGGGAATCTTTTTCAGAGTCTCT	799
Db	637	AGTCGATCTTTCAAGTACCCCGAGGACATGCCCAATGGGAATCTTTTTCAGAGTCTCT	696
Qy	800	CAGCTGCCC	809
Db	697	AACCTTGCCC	706
RESULT 8			
AAS30481			
ID	AAS30481 standard; DNA; 7453 BP.		
XX	AAS30481;		
XX	21-NOV-2001 (first entry)		
DT	DNA encoding novel prostate gland antigen, Seq ID No 339.		
DE	Human; nontropic; neuroprotective; cytostatic; antiparkinsonian;		
KW	antianemic; dermatological; immunosuppressive; antiinflammatory;		
KW	antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;		
KW	osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;		
KW	prostatosis; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;		
KW	hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;		
KW	reproductive system disorder; autoimmune disorder; urinary system;		
KW	systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;		
KW	blood-related disorder; hyperproliferative disorder; respiratory;		
KW	neurological disorder; endocrine disorder; inflammatory disorder;		
KW	liver disorder; wound healing; food preservative; ds.		
XX	Homo sapiens.		
OS	WO200155447-A1.		
PN	02-AUG-2001.		
XX	17-JAN-2001; 2001WO-US01330.		
PF	31-JAN-2000; 2000US-0179065.		
XX	04-FEB-2000; 2000US-0180628.		
PR	24-FEB-2000; 2000US-0184664.		
PR	02-MAR-2000; 2000US-0186350.		
PR	16-MAR-2000; 2000US-0189874.		
PR	17-MAR-2000; 2000US-0190076.		
PR	18-APR-2000; 2000US-0198123.		
PR	19-MAY-2000; 2000US-0205515.		
PR	07-JUN-2000; 2000US-0209467.		
PR	28-JUN-2000; 2000US-0214886.		
PR	30-JUN-2000; 2000US-0215135.		
PR	07-JUL-2000; 2000US-0216647.		
PR	07-JUL-2000; 2000US-0216880.		
PR	11-JUL-2000; 2000US-0217487.		
PR	11-JUL-2000; 2000US-0217496.		
PR	14-JUL-2000; 2000US-0218290.		
PR	26-JUL-2000; 2000US-0220963.		
PR	26-JUL-2000; 2000US-0220964.		
PR	14-AUG-2000; 2000US-0224518.		

CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),  
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
CC disease or atherosclerosis). The compositions and methods may also be  
CC used in assessing the histological type of neoplasm associated with  
CC ovarian cancer, monitoring the progression of ovarian cancer,  
CC determining whether ovarian cancer has metastasized or is likely to  
CC metastasize, selecting a composition for inhibiting ovarian cancer,  
CC assessing the ovarian carcinogenic potential of a compound, or  
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The  
CC present nucleic acid sequence encodes one of the ovarian cancer markers  
CC described in the invention.  
XX  
SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;  
  
Query Match 64.8%; Score 743; DB 24; Length 754;  
Best Local Similarity 99.9%; Pred. No. 1.4e-218;  
Matches 754; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
QY 301 GGAGTATGAGATGAACGAAATGCGAGAAATGAGTGGCGGTGAGTAAATGAGTTTCT 360  
Db |||||  
QY 361 GTCCAGCTCAAGATGACCTCAAGGAGGCAATGAATATCTATGTTGAGCGGATGCCA 420  
Db |||||  
QY 61 GTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGAATATCTATGTTGAGCGGATGCCA 120  
QY 421 AGGAAAGCATPAGGAGTTTCAAAATGACCGGAGACCTTAAGATGCCAGATCTGCTGA 480  
Db |||||  
QY 121 AGGAAAGCATPAGGAGTTTCAAAATGACCGGAGACCTTAAGATGCCAGATCTGCTGA 180  
QY 481 GTGTAATAGGCTGCATCTCTCTGAGGAGGAGAGATCTTTGGCGAGAGTCAAGCATGTTGGG 540  
Db |||||  
QY 181 GTGTAATAGGCTGCATCTCTCTGAGGAGGAGAGATCTTTGGCGAGATCAAGCATGTTGGG 240  
QY 541 CCTCAAGATCACCTACTTTGCACTGATGGATGAAAGGTTATGACATCACAGAGTGGGC 600  
Db |||||  
QY 241 CCTCAAGATCACCTACTTTGCACTGATGGATGAAAGGTTATGACATCACAGAGTGGGC 300  
QY 601 TGATGCCAGCTGTAGGTATCTCCCGAGATACCCAGAGTCCCTTATCAGATCTCAAT 660  
Db |||||  
QY 301 TGGATGCCAGCTGTAGGTATCTCCCGAGATACCCAGAGTCCCTTATCAGATCTCAAT 360  
QY 661 TGGTTCTCGGATTCAGGACACAGAGGCGGAGAGGCCACCCAGATGCCCTCTCTGC 720  
Db |||||  
QY 361 TGGTTCTCGGATTCAGGACACAGAGGCGGAGAGGCCACCCAGATGCCCTCTCTGC 420  
QY 721 TGATCTTCAGGATTTCTTGAGTCCGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG 780  
Db |||||  
QY 421 TGATCTTCAGGATTTCTTGAGTCCGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG 479  
QY 781 GAACCTTTCTTGAGCTCTCAGCTGCGCTGCGCTGAGCGGCTGCGAGCTCTAAGCCCAACAG 840  
Db |||||  
QY 480 GAACCTTTCTTGAGCTCTCAGCTGCGCTGCGCTGAGCGGCTGCGAGCTCTAAGCCCAACAG 539  
QY 841 CACAGTACCAAGGAGAGCCAAACCTAAGCGGCGGAGAAAGTGAAGAGGCGCTTCCA 900  
Db |||||  
QY 540 CACAGTACCAAGGAGAGCCAAACCTAAGCGGCGGAGAAAGTGAAGAGGCGCTTCCA 599  
QY 901 ACCTTCATGCCCTTCTCTTCTCTCAAAATCAATGTCAGGAGTCAAGAGGCGGTAGCAC 960  
Db |||||  
QY 600 ACCTTCATGCCCTTCTCTTCTCTCAAAATCAATGTCAGGAGTCAAGAGGCGGTAGCAC 659  
QY 961 AGGATGGAGTTTGAATTTATCCCTCTCTCCCGGACACCTAGGAACTGAAATCTTTTCTTTT 1020  
Db |||||  
QY 660 AGGATGGAGTTTGAATTTATCCCTCTCTCCCGGACACCTAGGAACTGAAATCTTTTCTTTT 719  
QY 1021 TATTTTTCAGATGGAGTCTTGCTCTGTGGCCAG 1055  
Db |||||  
QY 720 TATTTTTCAGATGGAGTCTTGCTCTGTGGCCAG 754

RESULT 7

AAI96802  
ID AAI96802 standard; cDNA; 747 BP.  
XX  
AC AAI96802;  
XX  
DT 13-NOV-2001 (first entry)  
XX  
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 2877.  
XX  
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200166719-A1.  
XX  
PD 13-SEP-2001.  
XX  
PF 02-MAR-2001; 2001WO-JP01629.  
XX  
PR 07-MAR-2000; 2000JP-0159195.  
XX  
PA (CHIB-) CHIBA PREFECTURE.  
XX  
PA (HISM) HISAMITSU PHARM CO LTD.  
XX  
PI Nakagawara A;  
XX  
PI WPI; 2001-565584/63.  
XX  
DR Nucleic acids originating in gene expressed in human neuroblastoma,  
XX  
PT useful as probe or primer in diagnosing prognosis of human  
XX  
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker  
XX  
PT for anti-cancer agents -  
XX  
PS Claim 1; Page 2103; 2979pp; Japanese.  
XX  
CC The invention relates to novel genes (AAI93926-AAI97963) expressed in  
XX  
CC human neuroblastoma. The nucleic acids are applicable as a probe or  
XX  
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and  
XX  
CC susceptibility indicators or tumour markers for anti-cancer agents. The  
XX  
CC gene information for diagnosing prognosis is related to factors similar  
XX  
CC to that for N-myc and TrkA genes.  
XX  
SQ Sequence 747 BP; 191 A; 167 C; 200 G; 165 T; 24 other;  
  
Query Match 54.8%; Score 628.2; DB 22; Length 747;  
Best Local Similarity 97.0%; Pred. No. 3.7e-183;  
Matches 650; Conservative 0; Mismatches 19; Indels 1; Gaps 1;  
  
QY 141 TGGGGTTTGGGCCACAGCATCAGATGTTGAACTGAAGAGGCTTAGACACTGGCAG 200  
Db |||||  
QY 37 TGGGCACCTTGGCCCTACTGGCAGATGTTGAACTGAAGAGGCTTAGACACTGGCAG 96  
QY 201 TGATGGTTTCATCTGACAAAATCATATCCCGGCTGAGGAGGCTTCAAGGTTTGC 260  
Db |||||  
QY 97 TGATGGTTTCATCTGACAAAATCATATCCCGGCTGAGGAGGCTTCAAGGTTTGC 156  
QY 261 GAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAAGAGTATGAGATCAACGAA 320  
Db |||||  
QY 157 GAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAAGAGTATGAGATCAACGAA 216  
QY 321 TGGCAGAGAAATGAGCTGAGCCGGTCAAGTAAATGAGTTTCTGTCCAAAGCTGCAAGATGACC 380  
Db |||||  
QY 217 TGGCAGAGAAATGAGCTGAGCCGGTCAAGTAAATGAGTTTCTGTCCAAAGCTGCAAGATGACC 276  
QY 381 TCAAGAGGCAATGAATATCTATGTTAGCCGATGCCAAGAAAGCATAGAGGTTTG 440  
Db |||||  
QY 277 TCAAGAGGCAATGAATATCTATGTTAGCCGATGCCAAGAAAGCATAGAGGTTTG 336  
QY 441 AAATGAGCGGGAACCTAAGAGTGCAGATCTGTGCTGAGTGAATAGGCTGCATCTTG 500  
Db |||||  
QY 337 AAATGAGCGGGAACCTAAGAGTGCAGATCTGTGCTGAGTGAATAGGCTGCATCTTG 396  
QY 501 CTGAGGAAGGAGACTTTTGGGCAGAGTCAAGCATGTGGGCGCTCAAGATCACTACTTTG 560

immunogen ), binding partners of the polypeptides, a method for identifying an agent useful in modulating mesenchymal cell differentiation induction activity of a molecule, a method of diagnosing a condition characterized by aberrant expression of a nucleic acid molecule or its expression product; a method for determining regression, progression or onset of cartilaginous tissue degeneration condition in a subject characterized by aberrant expression of a nucleic acid molecule or its expression product, a method for treating a cartilaginous tissue degeneration condition, a method for treating a subject to reduce the risk of cartilaginous tissue degeneration condition developing in the subject, a method for identifying a candidate agent for treating a cartilaginous tissue degeneration condition, and a solid-phase nucleic acid molecule array consisting essentially of a set of nucleic acid molecule as cited above (or known from known genes shown to be differentially expressed in developing mesenchymal cells using the technique of representational difference analysis, RDA), its expression products or fragments, fixed to a solid substrate. The nucleic acids, polypeptides and agents are useful for treating cartilaginous tissue degeneration conditions such as osteoarthritis, rheumatoid arthritis, gout arthritis, adjuvant arthritis, arthritis deformans, infectious arthritis or osteochondrosis. The present sequence is a cDNA from a known gene differentially expressed in developing mesenchymal cells.

XX  
SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;

Query Match 64.8%; Score 743; DB 24; Length 754;  
Best Local Similarity 99.9%; Pred. No. 1.4e-218;  
Matches 754; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 301 GGAGTATGAGTGAAGCAATGCGAGAGATGAGCTGAGCGGTGCTAGTAAATGAGTTTCT 360  
Db 1 GGAGTATGAGTGAAGCAATGCGAGAGATGAGCTGAGCGGTGCTAGTAAATGAGTTTCT 60

QY 361 GTCACAGCTGACAGTACCTCAAGGAGGCAATGACTATGATGTAGCCGATGCCA 420  
Db 61 GTCACAGCTGACAGTACCTCAAGGAGGCAATGACTATGATGTAGCCGATGCCA 120

QY 421 AGGAAAGCATAGGAGGTTTGAAATGACCGGAGCACTTAAGTGCCAGATACGTGCTGA 480  
Db 121 AGGAAAGCATAGGAGGTTTGAAATGACCGGAGCACTTAAGTGCCAGATACGTGCTGA 180

QY 481 GTGTAATAGCTGCATCTCTGCTGAGGAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG 540  
Db 181 GTGTAATAGCTGCATCTCTGCTGAGGAGGAGACTTTTGGCGAGAGTCAAGCATGTTGGG 240

QY 541 CTTCAAGATCACTACTTTGCACTGATGATGAAGAGTGTATGATCATCAGAGTGGGC 600  
Db 241 CTTCAAGATCACTACTTTGCACTGATGATGAAGAGTGTATGATCATCAGAGTGGGC 300

QY 601 TGGATCCAGCGTGTAGGTATCTCCCGAGATACCCAGAGTCCCGTATCATCTCAT 560  
Db 301 TGGATCCAGCGTGTAGGTATCTCCCGAGATACCCAGAGTCCCGTATCATCTCAT 360

QY 661 TGGTTCTCGATTCCAGGACAGAGGCGGAGAGGACCCAGATGCCCTCTCTGC 720  
Db 361 TGGTTCTCGATTCCAGGACAGAGGCGGAGAGGACCCAGATGCCCTCTCTGC 420

QY 721 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCAATGG 780  
Db 421 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCAATGG 479

QY 781 GAACTTCTTGAGCTCCCTGAGCTGCGCTGAGCGGTGAGCTCTTAAGCCCAAG 840  
Db 480 GAACTTCTTGAGCTCCCTGAGCTGCGCTGAGCGGTGAGCTCTTAAGCCCAAG 539

QY 841 CACAGTACCCAGGAGGAGCCAAACCTTAAGCGGGGAGAAAGTGAGGAGGCCCTTCCA 900  
Db 540 CACAGTACCCAGGAGGAGCCAAACCTTAAGCGGGGAGAAAGTGAGGAGGCCCTTCCA 599

QY 901 AGTTGATGCCCTCTCTTTCTCAATGTCAGGAGTCAAAAGGGGTGTAGCAC 960  
Db 600 AGTTGATGCCCTCTCTTTCTCAATGTCAGGAGTCAAAAGGGGTGTAGCAC 659

QY 961 AGGATGGAGTTTGATTTATCCCTCTCCCCCAACACCTAGGAACCTGAATCTTTTCTTTT 1020  
Db 660 AGGATGGAGTTTGATTTATCCCTCTCCCCCAACACCTAGGAACCTGAATCTTTTCTTTT 719

QY 1021 TATTTTTCGATGGAGTCTTCTCTCTGTTGCCCGAG 1055  
Db 720 TATTTTTCGATGGAGTCTTCTCTCTGTTGCCCGAG 754

RESULT 6  
ABS76460  
ID ABS76460 standard; cDNA; 754 BP.  
XX  
AC ABS76460;  
DT 11-DEC-2002 (first entry)  
XX  
DE cDNA encoding human ovarian cancer marker OV38.  
XX  
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;  
KW central nervous system disorder; bacterial meningitis; viral meningitis;  
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;  
KW brain herniation; inflammation; encephalitis; testicular disorder;  
KW nontuberculous granulomatous orchitis; connective tissue disorder;  
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
KW histological type; carcinogenic; ovarian cancer marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200271928-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 14-MAR-2002; 2002WO-US07826.  
XX  
PR 14-MAR-2001; 2001US-276025P.  
PR 14-MAR-2001; 2001US-276026P.  
PR 10-AUG-2001; 2001US-311732P.  
PR 19-SEP-2001; 2001US-323580P.  
PR 26-SEP-2001; 2001US-324967P.  
PR 26-SEP-2001; 2001US-325102P.  
PR 26-SEP-2001; 2001US-325149P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;  
PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Viefy PO, Mills GB;  
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K,  
XX  
XX WPI; 2002-723277/78.  
DR P-PSDB; ABG96364.  
DR  
PT Assessing whether a patient is afflicted with ovarian cancer, useful in  
PT assessing the stage or progression of the disease, comprises comparing  
PT the expression level of a cancer marker in a sample from a patient and  
PT from a non cancer patient -  
XX  
PS Disclosure; Page 300; 481pp; English.  
XX  
CC The present invention relates to a new method for assessing whether a  
CC patient is afflicted with ovarian cancer. The method involves comparing  
CC the expression level of a marker in a patient sample and the normal level  
CC of expression of the marker in a control non-ovarian cancer sample, where  
CC the marker is selected from 363 cancer markers described in the  
CC specification. The method of the invention is useful in diagnosing or  
CC characterising cancer, in detecting the presence of cancer as early as  
CC possible, and the recurrence of ovarian cancer. The method may also be of  
CC particular use with patients having an enhanced risk of developing  
CC ovarian cancer (e.g. patients having a familial history of ovarian  
CC cancer). The cancer markers may be used in the management and treatment  
CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),

PD 14-OCT-1999.  
XX 29-MAR-1999; 99WO-US06831.  
XX 03-APR-1998; 98US-0054956.  
XX (CURA-) CURAGEN CORP.  
XX Nandabalan K, Kingsmore S;  
XX WPI; 1999-620203/53.  
DR P-PSDB; AAY32126.  
XX  
PT Protein complexes, interacting proteins, and related polynucleotides  
PT useful for treating and preventing e.g. atopic, autoimmune or  
PT neurodegenerative diseases -  
XX  
PS Claim 21; Fig 7; 172pp; English.  
XX  
XX This is the nucleotide sequence of cDNA which codes for a novel  
CC human LYST interacting protein, LIP6 (see AAY32126), that shows  
CC homology to pestivirus Nsp-3. LYST is the human lysosomal  
CC Chediak-Higashi syndrome (CHS) protein. The invention relates to  
CC complexes of LYST or LYST-2 (see AAY32120) with proteins identified  
CC as interacting with LYST or LYST-2 by a modified yeast two-hybrid  
CC assay system. The interacting proteins include 10 novel proteins,  
CC LIP1-10 (see AAY32121-30). Methods of screening the protein complexes  
CC for efficacy in treating and/or preventing atopic diseases (e.g.  
CC asthma, nasal polyps, hay fever rhinitis, urticaria) autoimmune  
CC diseases (e.g. CHS, rheumatoid arthritis, systemic lupus  
CC erythematosus, inflammatory bowel disease, diabetes mellitus,  
CC multiple sclerosis), neurodegenerative disease, certain forms  
CC of cancer, pigmentation disorders, platelet dysfunction and viral  
CC diseases are provided. Nucleic acids (see AAZ34487-96) encoding  
CC LIP1-10, modulation of LIP function by gene therapy, use of  
CC antisense oligonucleotides for suppression of LIP protein  
CC expression, screening for agonists and antagonists, diagnosing or  
CC screening for the presence of a predisposition to a disease or  
CC disorder, and animal models are also disclosed.  
XX  
SQ Sequence 754 BP; 193 A; 185 C; 200 G; 176 T; 0 other;  
  
Query Match 54.8%; Score 743; DB 20; Length 754;  
Best Local Similarity 99.9%; Pred. No. 1.4e-218;  
Matches 754; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
QY 301 GGAGTATGAGATCAACGAGATGCAGAGATGAGCTGAGCGGCTCAGTAATCAGTTTCT 360  
DB 1 GGAGTATGAGATGAAACGAGATGCAGAGATGAGCTGAGCGGCTCAGTAATCAGTTTCT 60  
  
QY 361 GTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGACTATGATGTAGCCGATGCCA 420  
DB 61 GTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGACTATGATGTAGCCGATGCCA 120  
  
QY 421 AGGAAAGCATAGGAGGTTTGAATGACCGGGAACCTAAGAGTGCCAGATACCTGCTGA 480  
DB 121 AGGAAAGCATAGGAGGTTTGAATGACCGGGAACCTAAGAGTGCCAGATACCTGCTGA 180  
  
QY 481 GTGTATAGGCTGCATCCTGCTGAGGAGGAGAGACTTTGGGAGAGTCAAGCATGTGGG 540  
DB 181 GTGTATAGGCTGCATCCTGCTGAGGAGGAGAGACTTTGGGAGAGTCAAGCATGTGGG 240  
  
QY 541 CCTCAAGATCACCTACTTTTGCATGTGATGAAAGGTTGATGATCAGATCAGAGTGGGC 600  
DB 241 CCTCAAGATCACCTACTTTTGCATGTGATGAAAGGTTGATGATCAGATCAGAGTGGGC 300  
  
QY 601 TGGATGCCAGCGTGTAGGTATCTCCACAGATACCCACAGAGTCCCTATCAGATCTCAT 660  
DB 301 TGGATGCCAGCGTGTAGGTATCTCCACAGATACCCACAGAGTCCCTATCAGATCTCAT 360  
  
QY 661 TGGTTCTCGATTCCAGGACCCAGAGGGGGGAGAGGACCCAGATGCCCTCTCTGC 720  
DB 361 TGGTTCTCGATTCCAGGACCCAGAGGGGGGAGAGGACCCAGATGCCCTCTCTGC 420

QY 721 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATGCCAATGG 780  
DB 421 TGATCTTCAGGATTTCTTGAGTCGGATCTTTCAAGTACCCCGAGGCGAGATG-CCAATGG 479  
  
QY 781 GAACTTTCTTTGAGCTCCTCAGCTGCCCTCGAGCGGCTGAGCCCTCTAAGCCCAACAG 840  
DB 480 GAACTTTCTTTGAGCTCCTCAGCTGCCCTCGAGCGGCTGAGCCCTCTAAGCCCAACAG 539  
  
QY 841 CACAGTACCCAGGAGAGCCAAACCTAAGCGGCGGAGAAAGTAGAGGAGGCGCTTCCA 900  
DB 540 CACAGTACCCAGGAGAGCCAAACCTAAGCGGCGGAGAAAGTAGAGGAGGCGCTTCCA 599  
  
QY 901 ACCTGATGCCCTCTCTCTTCTCAATCAATGTGAGGAGTCAAAAGGCTGTAGCAC 960  
DB 600 ACCTGATGCCCTCTCTCTTCTCAATCAATGTGAGGAGTCAAAAGGCTGTAGCAC 659  
  
QY 961 AGGATGGAGTTTGATTTATCTCTCTCCCTCCCTCCCTAGCAACTGAATCTTTTCTTT 1020  
DB 660 AGGATGGAGTTTGATTTATCTCTCTCCCTCCCTCCCTAGCAACTGAATCTTTTCTTT 719  
  
QY 1021 TATTTTGGAGTGGAGTCTTGCTCTGTTGCCCGAG 1055  
DB 720 TATTTTGGAGTGGAGTCTTGCTCTGTTGCCCGAG 754

RESULT 5  
ABX04184  
ID ABX04184 standard; cDNA; 754 BP.  
XX ABX04184;  
AC ABX04184;  
DT 10-JAN-2003 (first entry)  
DE Human mRNA differentially expressed in mesenchymal cells #31.  
XX  
KW Human; ss; gene; skeletal growth; cartilage degeneration disorder;  
KW chondroblastic phenotype; mesenchymal cell; cartilage formation;  
KW bone formation; arthritis; osteoarthritis; rheumatoid arthritis;  
KW gout arthritis; adjuvant arthritis; arthritis deformans; antigout;  
KW infectious arthritis; osteochondrosis; RDA; antiarthritic; osteopathic;  
KW antirheumatic; antiinflammatory; representational difference analysis.  
XX  
OS Homo sapiens.  
XX  
PN WO200271927-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 12-MAR-2002; 2002WO-US07787.  
XX  
PR 12-MAR-2001; 2001US-274980P.  
XX  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
XX  
PI Yates KE, Mizuno S, Glowacki J;  
XX  
XX WPI; 2002-723276/78.  
XX  
PT New nucleic acid molecules capable of promoting chondrogenesis, useful  
PT for diagnosing and treating cartilaginous tissue degeneration  
PT conditions, e.g. osteoarthritis, rheumatoid arthritis, gout arthritis,  
PT or osteochondrosis -  
XX  
PS Claim 33; Page 129; 153pp; English.  
XX  
CC The invention relates to new isolated nucleic acid molecule comprising a  
CC nucleic acid molecule consisting of a gene differentially expressed  
CC in cells undergoing differentiation from mesenchymal cell to a  
CC chondroblastic phenotype, or hybridising under stringent conditions  
CC to them (or their fragments). Also included are expression vectors,  
CC transformed host cells, expressed polypeptides or peptide fragments  
CC (which induce differentiation of a mesenchymal cell and may be as an

OS	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	FT	122-2230
XX	FT	/*tag=
XX	FT	/product= "P125-77.22 polypeptide"
XX	PN	WO200226810-A1.
XX	PD	04-APR-2002.
XX	PF	10-SEP-2001; 2001WO-CN01354.
XX	PR	12-SEP-2000; 2000CN-0125190.
XX	PA	(SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX	PI	Mao Y, Xie Y;
XX	WPI	2002-281319/32.
XX	DR	P-PSDB; ABB77732.
XX	CC	Protein P125-77.22 and encoding polynucleotide, used in diagnosis and treatment of human mucosal disease caused by BVDV infection -
XX	CC	The polynucleotide may also be used for gene therapy.
XX	PS	Claim 6; Page 27-29; 33pp; Chinese.
XX	CC	The present sequence encodes human P125-77.22 polypeptide. The polypeptide and polynucleotide are used in diagnosis and treatment of human mucosal disease caused by viral BVDV (undefined) infection.
XX	CC	The polynucleotide may also be used for gene therapy.
XX	SQ	Sequence 3286 BP; 787 A; 785 C; 895 G; 819 T; 0 other;
XX	Query Match	99.6%; Score 1142.2; DB 24; Length 3286;
XX	Best Local Similarity	99.7%; Pred. No. 0;
XX	Matches 1144;	Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1	GCCTGGGTCAACGAGATATTAAATAGCGAGGGGAATGCACCTGTAGCTAGTGGCGCTA 60
DB	1324	GCCTGGGTCAACGAGATATTAAATAGCGAGGGGAATGCACCTGTAGCTAGTGGCGCTA 1383
QY	61	CTGCCAGCCTGAAGAGAGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCAGGATGA 120
DB	1384	CTGCCAGCCTGAAGAGAGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCAGGATGA 1443
QY	121	GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAATGAAGAA 180
DB	1444	GCTAAACCCCTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTGAATGAAGAA 1503
QY	181	GGCCTATAGACAGCTGGCAGTGTGTTCTATCTGACAAATCATCATCCCGGCTGA 240
DB	1504	GGCCTATAGACAGCTGGCAGTGTGTTCTATCTGACAAATCATCATCCCGGCTGA 1563
QY	241	GGAGGCTTCAAGGTTTTCGGACGCTGGGACATTGGGACATTGACAAATGCTGAAAGGAAA 300
DB	1564	GGAGGCTTCAAGGTTTTCGGACGCTGGGACATTGACAAATGCTGAAAGGCAA 1623
QY	301	GGAGTATGAGATGAAACGAAATGCAGAGATGAGCTGAGCGGTGAGTAAATGAGTTCT 360
DB	1624	GGAGTATGAGATGAAACGAAATGCAGAGATGAGCTGAGCGGTGAGTAAATGAGTTCT 1683
QY	361	GTCCAAAGCTGCAAGATGAGCTCAAGAGGCAATGAATATGATGCTGAGCGGATGCCA 420
DB	1684	GTCCAAAGCTGCAAGATGAGCTCAAGAGGCAATGAATATGATGCTGAGCGGATGCCA 1743
QY	421	AGGAAGCATAGGAGTTTGAATGACCGGGAACCTAAGATGCCAGATCTGTGCTGA 480
DB	1744	AGGAAGCATAGGAGTTTGAATGACCGGGAACCTAAGATGCCAGATCTGTGCTGA 1803
QY	481	GTGTAATAGGCTGCATCCCTGCTGAGGAAGAGACTTTTGGGAGAGTCAAGCATGTTGG 540
DB	1804	GTGTAATAGGCTGCATCCCTGCTGAGGAAGAGACTTTTGGGAGAGTCAAGCATGTTGG 1863
QY	541	CCTCAAGATCACCTACTTTTGCACTGATGATGAAAAGGTGTATGATCATCACAGTGGGC 600
DB	1864	CCTCAAGATCACCTACTTTTGCACTGATGATGAAAAGGTGTATGATCATCACAGTGGGC 1923
QY	601	TGATGCCAGCGTGTAGGTATCTCCCCAGATACCCACAGAGTCCCCCTATCATCTCATT 660
DB	1924	TGATGCCAGCGTGTAGGTATCTCCCCAGATACCCACAGAGTCCCCCTATCATCTCATT 1983
QY	661	TGGTCTCGGATTCAGGACACAGAGGGGGGAGAGAGCCACCCAGATGCCCTCTCTGC 720
DB	1984	TGGTCTCGGATTCAGGACACAGAGGGGGGAGAGAGCCACCCAGATGCCCTCTCTGC 2043
QY	721	TGATCTTCAGGATTTCTTCAGTCGGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG 780
DB	2044	TGATCTTCAGGATTTCTTCAGTCGGATCTTTCAAGTACCCCGAGGAGATGCCCAATGG 2103
QY	781	GAACTTTCTTTTGCACTCTCAGCTCCCTGAGCGCTGAGCGCTGAGCGCTCTAAGGCCCAACAG 840
DB	2104	GAACTTTCTTTTGCACTCTCAGCTCCCTGAGCGCTGAGCGCTCTAAGGCCCAACAG 2163
QY	841	CACAGTACCCAAAGGGAGAGCAACCTAAGCGGGCGAAGAAAGTGAAGAGGCCCTTCCA 900
DB	2164	CACAGTACCCAAAGGGAGAGCAACCTAAGCGGGCGAAGAAAGTGAAGAGGCCCTTCCA 2223
QY	901	AGTTGATGCCCTTCTCTTCTCTCAATCAATGTGAGGAGTCAAAAGGGCTGTAGCAC 960
DB	2224	AGTTGATGCCCTTCTCTTCTCTCAATCAATGTGAGGAGTCAAAAGGGCTGTAGCAC 2283
QY	961	AGGATGGAGTTGATTTATCCCTCCCTCCCAACACCTAGGAACCTCAATCTTTTCTTTT 1020
DB	2284	AGGATGGAGTTGATTTATCCCTCCCTCCCAACACCTAGGAACCTCAATCTTTTCTTTT 2343
QY	1021	TATTTTTCAGATGAGTCTTCTCTCTGTTGCCAGCTGGAGTGCAGTGGTGTATCTCAG 1080
DB	2344	TATTTTTCAGATGAGTCTTCTCTCTGTTGCCAGCTGGAGTGCAGTGGTGTATCTCAG 2403
QY	1081	CTTACTGCAACCTGTCTCTCCGGGTTCAAGCAATTTCTCCATCTCAGGCTCTCTGAGTAG 1140
DB	2404	CTTACTGCAACCTGTCTCTCCGGGTTCAAGCAATTTCTCCATCTCAGGCTCTCTGAGTAG 2463
QY	1141	CTGGGAT 1147
DB	2464	CTGGGAT 2470
XX	RESULT 4	
XX	AZ34492	
XX	ID	AZ34492 standard; cDNA; 754 BP.
XX	AC	AZ34492;
XX	XX	
XX	DT	01-FEB-2000 (first entry)
XX	DE	Human LYST interacting protein LIP6 cDNA.
XX	XX	
XX	KW	LIP6; human; LYST-2; LYST interacting protein; lysosome; CHS;
XX	KW	Chediak-Higashi syndrome; autoimmune disease; rheumatoid arthritis;
XX	KW	systemic lupus erythematosus; inflammatory bowel disease;
XX	KW	diabetes mellitus; multiple sclerosis; atopic disease; asthma;
XX	KW	hay fever; rhinitis; urticaria; nasal polyp; cancer;
XX	KW	neurodegenerative disease; pigmentation disorder; viral disease;
XX	XX	platelet dysfunction; ss.
XX	OS	Homo sapiens.
XX	XX	
XX	FT	Key
XX	FT	Location/Qualifiers
XX	FT	11..586
XX	FT	/*tag= a
XX	FT	/partial
XX	PN	WO9951741-A2.
XX	XX	





```
PR 03-AUG-1999; 99US-0146908.
XX 22-OCT-1999; 99US-0160924.
PA (INCY-) INCYTE GENOMICS INC.
XX Yue H, Bandman O, Tang YT, Baughn MR, Azimzai Y, Lu DAM;
XX WPI; 2001-159853/16.
DR P-PSDB; AAB67455.
XX
PT New human chaperone proteins and polynucleotides, useful in diagnosing,
FT treating and preventing reproductive, eye, neuromuscular, metabolic,
PT autoimmune or inflammatory disorders
XX
XX Claim 5; Page 95-96; 102pp; English.
XX
CC The present sequence encodes a human chaperone polypeptide. Human
CC chaperone polypeptides and polynucleotides are useful in the diagnosis,
CC treatment and prevention of reproductive (e.g. prolactin production,
CC infertility, endometrial or ovarian tumour, cancer of the breast,
CC prostate or testis, Peyronie's disease), eye (e.g. conjunctivitis,
CC keratitis, iritis, glaucoma), neuromuscular, metabolic (e.g. Zellweger
CC syndrome, Addison's disease, cystic fibrosis), and autoimmune and
CC inflammatory disorders (e.g. systemic lupus erythematosus, acquired
CC immunodeficiency syndrome (AIDS), asthma, atherosclerosis), infectious
CC or viral diseases, and cell proliferative disorders. Chaperone
CC polynucleotides may be used for somatic or germline gene therapy, to
CC detect and quantify gene expression in biopsied tissues in which
CC expression is correlated with disease.
XX
XX Sequence 1147 BP; 287 A; 278 C; 314 G; 268 T; 0 other;

Query Match 100.0%; Score 1147; DB 22; Length 1147;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTGGGTCAAGCAGAAATATTATAGCGGGGAATGCACCTGTAGTGGGGCTA 60
DB 1 GCCTTGGGTCAAGCAGAAATATTATAGCGGGGAATGCACCTGTAGTGGGGCTA 60

QY 61 CTCCAGCCTGAAGAGAAAGTGGCTCGACTTTGACCATGGCTGGGGTTCCTGAGGATGA 120
DB 61 CTCCAGCCTGAAGAGAAAGTGGCTCGACTTTGACCATGGCTGGGGTTCCTGAGGATGA 120

QY 121 GCTAAACCTTTTCATGTACTGGGGTTGAGGCGACAGCATCAGATGTGAATGAAGAA 180
DB 121 GCTAAACCTTTTCATGTACTGGGGTTGAGGCGACAGCATCAGATGTGAATGAAGAA 180

QY 181 GGCTTATAGACAGCTGGCAGTGTGTTTCATCTCGACAAATCATCTCCCGGCTGA 240
DB 181 GGCTTATAGACAGCTGGCAGTGTGTTTCATCTCGACAAATCATCTCCCGGCTGA 240

QY 241 GGAGGCTTCAAGGTTTTCGAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAA 300
DB 241 GGAGGCTTCAAGGTTTTCGAGCAGCTTGGGACATTTGTCAGCAATGCTGAAAAGCGAAA 300

QY 301 GGAGTATGAGATGAACGAAATGCGAGAAATGAGTGGCGGTGAGTAAATGAGTTCT 360
DB 301 GGAGTATGAGATGAACGAAATGCGAGAAATGAGTGGCGGTGAGTAAATGAGTTCT 360

QY 361 GTCCAGCTGCAAGATGACCTCAAGAGGCAATGATATGATGTAGCCGATGCCA 420
DB 361 GTCCAGCTGCAAGATGACCTCAAGAGGCAATGATATGATGTAGCCGATGCCA 420

QY 421 AGGAAAGCATAGGAGTTTGAATGTGACCGGGAACCTTAAGAGTGCAGATCTGTCTGA 480
DB 421 AGGAAAGCATAGGAGTTTGAATGTGACCGGGAACCTTAAGAGTGCAGATCTGTCTGA 480

QY 481 GTGTAATAGCTGCATCTCTGCTAGGAGGAGACATTTTGGGAGAGTCAAGCATGTGGG 540
DB 481 GTGTAATAGCTGCATCTCTGCTAGGAGGAGACATTTTGGGAGAGTCAAGCATGTGGG 540

QY 541 CCTCAAGATCACTACTTTTGCACTGATGATGGAAGGTTGTATGATACACAGATGGGC 600
DB 541 CCTCAAGATCACTACTTTTGCACTGATGATGGAAGGTTGTATGATACACAGATGGGC 600

QY 601 TGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCCTATCAGATCTCAT 660
DB 601 TGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCCTATCAGATCTCAT 660

QY 661 TGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCCTATCAGATCTCAT 660
DB 661 TGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTCCCCTATCAGATCTCAT 660

QY 720 TGGTTCCTCGGATTCAGGACACAGAGCGGCGGAGAGGACCCACAGATGCCCTCCTGCG 720
DB 720 TGGTTCCTCGGATTCAGGACACAGAGCGGCGGAGAGGACCCACAGATGCCCTCCTGCG 720

QY 780 TGAATCTTCAGGATTTCTTGAATCGGATCTTTCAAGTACCCCGGAGGAGATGCCCAATGG 780
DB 780 TGAATCTTCAGGATTTCTTGAATCGGATCTTTCAAGTACCCCGGAGGAGATGCCCAATGG 780

QY 840 GAACTTCTTTGAGCTCCTCAGCTCCCTCGAGCGCTGGAGCGGCTTAAAGCCCAACAG 840
DB 840 GAACTTCTTTGAGCTCCTCAGCTCCCTCGAGCGGCTTAAAGCCCAACAG 840

QY 900 CACAGTACCCAGGAGAGCCAAACCTTAAGCGGCGGAGAAAGTGAAGAGGCGCTTCCA 900
DB 900 CACAGTACCCAGGAGAGCCAAACCTTAAGCGGCGGAGAAAGTGAAGAGGCGCTTCCA 900

QY 960 ACCTTTCAGTCCCTCTCTTCTCCTCAAAATCAATGTCAGGAGTCAAAAGGCGTGTAGCAC 960
DB 960 ACCTTTCAGTCCCTCTCTTCTCCTCAAAATCAATGTCAGGAGTCAAAAGGCGTGTAGCAC 960

QY 1020 AGGATGGAGTTTGAATTTATCCCTCCCTCCCGGCTTCAAGCAATCTCTTCTTTT 1020
DB 1020 AGGATGGAGTTTGAATTTATCCCTCCCTCCCGGCTTCAAGCAATCTCTTCTTTT 1020

QY 1080 TATTTTTCAGATGGAGTCTTCTGTTGCCAGCTGGAGTGCAGTGTGATCTCAG 1080
DB 1080 TATTTTTCAGATGGAGTCTTCTGTTGCCAGCTGGAGTGCAGTGTGATCTCAG 1080

QY 1140 CTCTACTGCAACCTCTGCTCCCGGTTCAAGCAATCTCTCAGCTCTCAGCTCTCAGTAG 1140
DB 1140 CTCTACTGCAACCTCTGCTCCCGGTTCAAGCAATCTCTCAGCTCTCAGCTCTCAGTAG 1140

QY 1141 CTGGGAT 1147
DB 1141 CTGGGAT 1147

RESULT 2
ABL55088
ID ABL55088 standard; DNA; 2279 BP.
XX ABL55088;
AC ABL55088;
XX
DT 08-OCT-2002 (first entry)
XX
XX Human cDNA sequence #1 from clone HNTPB82.
XX
XX Human; HNTPB82; secreted protein; immunosuppressive; food preservative;
XX antihistatic; antirheumatic; antiproliferative; cytostatic; cardiac;
XX vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
XX virucide; fungicide; ophthalmological; vulnary; gene therapy; ELISA;
XX radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;
XX hyperproliferative disorder; cardiovascular disorder; angiogenesis;
XX cerebrovascular disorder; nervous system disorder; ocular disorder;
XX wound healing; food additive; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 36..1274
XX FT /*tag= a
XX FT sig_peptide 36..38
XX FT /*tag= b
XX FT mat_peptide 39..1271
XX FT /*tag= c
XX
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2003, 07:30:31 ; Search time 379 Seconds  
(without alignments)

8169.539 Million cell updates/sec

Title: US-10-049-742-22

Perfect score: 1147

Sequence: 1 gccttggtggaagcagaata.....gcctctgagtagctggat 1147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1147	100.0	1147	22	AAF54994
2	1145.4	99.9	2279	24	ABL55088
3	1142.2	99.6	3286	24	ABL56700
4	743	64.8	754	20	AAZ34492
5	743	64.8	754	24	ABX04184
6	743	64.8	754	24	ABX04184
7	628.2	54.8	747	22	AAZ96802
8	453	39.5	7453	22	AAU07447

9	453	39.5	7453	22	AAU06261	Human reproductive
10	272.8	23.8	884	24	ABL55114	Human cDNA sequenc
11	251.4	21.9	12119	22	AAC86936	Nucleotide sequenc
12	251.4	21.9	12578	21	AAZ36196	Nucleotide sequenc
13	251.4	21.9	12734	24	ABA95615	Chimeric BVDV/HCV
14	251.4	21.9	12842	21	AAZ36203	Nucleotide sequenc
15	251.4	21.9	13198	21	AAZ36211	Nucleotide sequenc
16	251.4	21.9	14078	21	AAA38807	Plasmid pBVdDNI, c
17	251.4	21.9	14578	21	AAZ48136	Bovine viral diarr
18	251.4	21.9	15065	21	AAZ36195	Nucleotide sequenc
19	251.4	21.9	16622	21	AAZ36212	Nucleotide sequenc
20	245.2	21.4	256	24	ABQ56856	Human colon cancer
21	239.4	20.9	11674	21	AAZ36210	Nucleotide sequenc
22	234.6	20.5	339	24	ABL85522	Human ovarian canc
23	145	12.6	2277	22	AAK68770	Human immune/haema
24	120	10.5	565	23	ABV60340	Human prostate exp
25	119	10.4	330	16	AAZ4617	Human gene signatu
26	116.8	10.2	28001	24	ABS54410	Human hydroxymethy
27	116.8	10.2	28001	25	ABX93300	Gene encoding huma
28	116.4	10.1	3135	23	ABL25653	Drosophila melanog
29	115.6	10.1	1811	22	AAH14407	Human cDNA sequenc
30	115.6	10.1	3132	22	AAH18382	Human cDNA sequenc
31	113.6	9.9	498	24	ABK45936	cDNA encoding colo
32	113.2	9.9	4941	22	AAI98901	Human kidney relat
33	113.2	9.9	4941	22	AAI63251	Androgen receptor
34	113.2	9.9	160820	25	ABO76673	Genomic sequence #
35	112.8	9.8	2147	23	ABK42851	Human nervous syst
36	112.8	9.8	32174	22	ABA15665	Human nervous syst
37	112.8	9.8	32174	22	ABA19477	Human nervous syst
38	112.8	9.8	32174	22	ABA20359	Human nervous syst
39	112.8	9.8	32174	22	ABA21505	Human nervous syst
40	112.8	9.8	32174	22	AAI36280	Human musculoskele
41	112.8	9.8	32174	22	AAS32855	Human genomic DNA
42	112.8	9.8	32174	22	AAS34394	Human DNA for a no
43	112.8	9.8	32174	22	AAS30638	DNA encoding novel
44	112.8	9.8	32174	22	AAU03792	Human reproductive
45	112.8	9.8	32174	22	AAU07447	Human reproductive

ALIGNMENTS

RESULT 1

AAF54994

ID AAF54994 standard; DNA; 1147 BP.

XX AAF54994;

DT 15-MAY-2001 (first entry)

XX Nucleotide sequence of a human chaperone polypeptide.

DE Human; chaperone polypeptide; reproductive disease; prolactin production;  
DE infertility; tumour; cancer; Peyronie's disease; eye disorder; Glaucoma;  
DE conjunctivitis; keratitis; neuromuscular disorder; cystic fibrosis;  
DE metabolic disorder; Zellweger syndrome; Addison's disease; Iritis;  
DE autoimmune disorder; inflammatory disorder; systemic lupus erythematosus;  
DE acquired immunodeficiency syndrome; AIDS; asthma; atherosclerosis;  
DE cell proliferative disorder; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 98..907

FT /tag= a /product= "chaperone protein"

XX WO200109178-A2.

XX 08-FEB-2001.

XX 03-AUG-2000; 2000WO-US21313.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2003, 07:30:31 ; Search time 90 Seconds

(without alignments)  
5625.183 Million cell updates/sec

Title: US-10-049-742-22

Perfect score: 1147  
Sequence: 1 gccttgggtcaagcagaata.....gcctcctgagtagctggat 1147

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTOTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	251.4	21.9	12734	4	US-09-344-456-1
2	251.4	21.9	14078	3	US-09-433-262-1
3	251.4	21.9	14078	4	US-09-702-330-1
4	251.4	21.9	14578	3	US-08-859-694-1
C 5	116.8	10.2	28001	4	US-09-819-993-3
6	112.4	9.8	17493	4	US-09-804-471A-3
C 7	110.4	9.6	43069	4	US-09-292-542A-1
8	109.8	9.6	74962	4	US-09-685-853A-3
C 9	107	9.3	955	4	US-09-641-638-4
C 10	107	9.3	955	4	US-09-641-638-5
C 11	107	9.3	955	4	US-09-641-638-6
C 12	107	9.3	955	4	US-09-641-638-7
C 13	107	9.3	1000	4	US-09-641-638-650
C 14	106	9.2	2871	4	US-09-489-847-111
C 15	106	9.2	2876	4	US-09-489-847-22
C 16	106	9.2	43950	4	US-09-735-934A-3
C 17	106	9.2	43950	4	US-10-060-332-3
C 18	105	9.2	17493	4	US-09-804-471A-3
C 19	105.2	9.2	31208	4	US-09-852-067-3
C 20	104.8	9.1	2839	3	US-08-468-856A-5
C 21	104.8	9.1	2839	3	US-08-468-859A-5
22	104.6	9.1	562	3	US-09-385-982-442
23	104.6	9.1	606	3	US-09-385-982-413
24	104.6	9.1	632	3	US-09-385-982-177
25	104.6	9.1	657	3	US-09-385-982-335
26	104.4	9.1	653	3	US-09-385-982-324
C 27	104.4	9.1	59065	4	US-09-813-817-3

ALIGNMENTS

RESULT 1

US-09-344-456-1  
; Sequence 1, Application US/093444456A  
; Patent No. 6326137  
; GENERAL INFORMATION:  
; APPLICANT: Hong, Zhi  
; APPLICANT: Lai, Vicky C.H.  
; APPLICANT: Lau, Johnson Y.N.  
; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE-DEPENDENT CHIMERIC  
; TITLE OF INVENTION: PESTIVIRUS  
; FILE REFERENCE: IN01038  
; CURRENT APPLICATION NUMBER: US/09/344,456A  
; CURRENT FILING DATE: 1999-06-25  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 12734  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric  
; OTHER INFORMATION: Pestivirus  
US-09-344-456-1

C 28	104.4	9.1	59065	4	US-09-978-197-3	Sequence 3, Appli
C 29	104.2	9.1	72604	4	US-09-268-992-7	Sequence 7, Appli
C 30	104.2	9.1	72604	4	US-09-657-474-7	Sequence 7, Appli
C 31	104	9.1	3507	1	US-08-832-883-67	Sequence 67, Appli
32	104	9.1	3507	2	US-08-832-877-67	Sequence 67, Appli
33	104	9.1	162450	4	US-09-345-882-1	Sequence 1, Appli
34	103.8	9.0	63588	4	US-09-873-404-3	Sequence 3, Appli
C 35	103	9.0	111282	4	US-09-754-250-3	Sequence 3, Appli
36	102.8	9.0	1002	4	US-09-641-638-578	Sequence 578, App
C 37	102.8	9.0	1184	4	US-09-489-847-76	Sequence 76, Appli
38	102.8	9.0	5789	3	US-09-242-348-3	Sequence 3, Appli
39	102.8	9.0	26664	4	US-09-564-805-28	Sequence 28, Appli
40	102.4	8.9	62804	4	US-09-800-960-3	Sequence 3, Appli
41	102.4	8.9	80246	3	US-09-078-294-4	Sequence 4, Appli
42	102.4	8.9	80595	3	US-09-078-294-3	Sequence 3, Appli
C 43	102.4	8.9	148567	4	US-09-801-876B-3	Sequence 3, Appli
44	102.2	8.9	3001	4	US-09-539-3330-187	Sequence 187, App
C 45	102.2	8.9	18073	3	US-09-078-294-12	Sequence 12, Appli

QY	404	ATGTGTAGCGGATGCCAAGGAAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT	463
Db	5150	ATGTGAGCGGATGCCAAGGAAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGT	5209
QY	464	GCAGATACCTGCTGAGTGTATAGGCTGCTGCTGAGGAAGGAGACTTTTGGGCA	523
Db	5210	GCCAGATACCTGCTGAGTGTATAGGCTGCTGCTGAGGAAGGAGTCTTTTGGGCA	5269
QY	524	GAGTCAAGCATGTTGGGCTTCAAGATCACTATTTGCACTGATGATGAAAGGTTAT	583
Db	5270	GAGTCAAGCATGTTGGGCTTCAAAATCACTATTTGCGCTGATGATGAAAGGTTAT	5329
QY	584	GACATCACAGGTGGGCTGGATCCGAGCGTGTAGGTATCTCCCCAGATACCCAGAGTC	643
Db	5330	GATATCACAGGTGGGCTGGATCCGAGCGTGTGGAAATCTCCCCAGATACCCAGAGTC	5389
QY	644	CCCTATCATCTCTATTTGGTTCTCGGATTCCAGGCA	680
Db	5390	CCTTGTACATCTCTATTTGGTTCTCGGATTCCAGGATGCTTTCA	5426

RESULT 2

```
US-09-433-262-1
; Sequence 1, Application US/09433262
; Patent No. 6168942
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Sheppard, Mike
; TITLE OF INVENTION: ATTENUATED FORMS OF BOVINE VIRAL DIARRHEA VIRUS
; FILE REFERENCE: PC10435A
; CURRENT APPLICATION NUMBER: US/09/433,262
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 60/107,908
; EARLIER FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14078
; TYPE: DNA
; ORGANISM: Bovine Viral Diarrhea Virus
US-09-433-262-1

Query Match      21.9%; Score 251.4; DB 3; Length 14078;
Best Local Similarity 94.2%; Pred. No. 1.6e-68;
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 404 ATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 463
Db 4532 ATGTGCGAGCCGATGCCAGGGAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 4591
Qy 464 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGACTTTTGGGCA 523
Db 4592 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGTGTATTTGGGCA 4651
Qy 524 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTAT 583
Db 4652 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCGCTGATGGATGGAAGGTGTAT 4711
Qy 584 GACATCACAGATGGGCTGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTC 643
Db 4712 GATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCCGAGATACCCACAGAGTC 4771
Qy 644 CCCTATCACATCTCATTTGGTTCTCGGATTCAGGCA 680
Db 4772 CCTTGTACATCTCATTTGGTTTACGGATGCCCTTTCA 4808

RESULT 3
US-09-702-330-1
; Sequence 1, Application US/09702330
; Patent No. 6410032
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Sheppard, Mike
; TITLE OF INVENTION: ATTENUATED FORMS OF BOVINE VIRAL DIARRHEA VIRUS
; FILE REFERENCE: PC10435A
; CURRENT APPLICATION NUMBER: US/09/702,330
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/433,262
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14078
; TYPE: DNA
; ORGANISM: Bovine Viral Diarrhea Virus
US-09-702-330-1

Query Match      21.9%; Score 251.4; DB 4; Length 14078;
Best Local Similarity 94.2%; Pred. No. 1.6e-68;
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 404 ATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 463
Db 4532 ATGTGCGAGCCGATGCCAGGGAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 4591
```

```
Qy 464 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGACTTTTGGGCA 523
Db 4592 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGTGTATTTGGGCA 4651
Qy 524 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTAT 583
Db 4652 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCGCTGATGGATGGAAGGTGTAT 4711
Qy 584 GACATCACAGATGGGCTGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTC 643
Db 4712 GATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCCGAGATACCCACAGAGTC 4771
Qy 644 CCCTATCACATCTCATTTGGTTCTCGGATTCAGGCA 680
Db 4772 CCTTGTACATCTCATTTGGTTTACGGATGCCCTTTCA 4808

RESULT 4
US-08-859-694-1
; Sequence 1, Application US/08859694A
; Patent No. 6001613
; GENERAL INFORMATION:
; APPLICANT: Donis, Ruben O.
; APPLICANT: Vassiliev, Ventsislav B.
; TITLE OF INVENTION: A plasmid bearing a cDNA copy of the genome of bovine
; TITLE OF INVENTION: viral diarrhea virus, chimeric derivations thereof, and
; TITLE OF INVENTION: method of producing an infectious bovine viral diarrhea
; FILE REFERENCE: UNVNS1110
; CURRENT APPLICATION NUMBER: US/08/859,694A
; CURRENT FILING DATE: 1997-05-21
; EARLIER APPLICATION NUMBER: 60/018,246
; EARLIER FILING DATE: 1996-05-24
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14578
; TYPE: DNA
; ORGANISM: bovine viral diarrhea virus
US-08-859-694-1

Query Match      21.9%; Score 251.4; DB 3; Length 14578;
Best Local Similarity 94.2%; Pred. No. 1.6e-68;
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 404 ATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 463
Db 5032 ATGTGCGAGCCGATGCCAGGGAAGCATAGGAGGTTTGAAATGGACCGGAAACCTTAAGAGT 5091
Qy 464 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGACTTTTGGGCA 523
Db 5092 GCCAGATACCTGTGCTGAGTGTAAATAGGCTGCATCCTCTGAGGAAGAGTGTATTTGGGCA 5151
Qy 524 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCACTGATGGATGGAAGGTGTAT 583
Db 5152 GAGTCAGCATGTGGGCGCTCAAGATCACCTACTTTGCGCTGATGGATGGAAGGTGTAT 5211
Qy 584 GACATCACAGATGGGCTGGATGCCAGCGTGTAGGTATCTCCCGAGATACCCACAGAGTC 643
Db 5212 GATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCCGAGATACCCACAGAGTC 5271
Qy 644 CCCTATCACATCTCATTTGGTTCTCGGATTCAGGCA 680
Db 5272 CCTTGTACATCTCATTTGGTTTACGGATGCCCTTTCA 5308

RESULT 5
US-09-819-993-3/c
; Sequence 3, Application US/09819993
; Patent No. 6436692
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
```

```
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; FILE REFERENCE: CL001195
; CURRENT APPLICATION NUMBER: US/09/819,993
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28001
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)..(28001)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-993-3

Query Match      10.2%; Score 116.8; DB 4; Length 28001;
Best Local Similarity 88.2%; Pred. No. 6.9e-26;
Matches 127; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1004 CTGAATCTTTTCTTTTATTTTGTAGATGGAGTCTTGTCTGTTCGCCAGCTGGAGTG 1063
DB 23843 CTTTTTTTTTTTTTTTTTTTGTAGATGGAATCTTGTCTGTTCGCCAGCTGGAGTG 23784

QY 1064 CAGTGGTGTGATCTCAGCTTACTGCAACTCTGTCTCCGGTTCGAAGCAATTCCTCCCAT 1123
DB 23783 CAGTGGTGTGATCTCAGCTTACTGCAACTCTGTCTCCGGTTCGAAGCAATTCCTCTGA 23724

QY 1124 CTCAGCCTCCTGAGTAGTGGGAT 1147
DB 23723 CTCAGCCTCCTGAGTAGTGGGAT 23700

RESULT 6
US-09-804-471A-3
; Sequence 3, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)..(174493)
; OTHER INFORMATION: n = A,T,C or G
US-09-804-471A-3

Query Match      9.8%; Score 112.4; DB 4; Length 174493;
Best Local Similarity 88.4%; Pred. No. 4.9e-24;
Matches 122; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1009 TCTTTTCTTTTATTTTGTAGATGGAGTCTTGTCTGTTCGCCAGCTGGAGTGCAGTG 1068
DB 119174 TTTTCTTTTATTTTGTAGAGGAGTCTCGTCTGTTCGCCAGGCTGGAGTGCAGTG 119233

QY 1069 GTGTGATCTCAGTCTACTGCAACCTCTGTCTCCCGGTTCAAGCAATTCCTCCATCTCAG 1128
DB 119234 GTGTGATCTCAGTCTACTGCAACCTCTGTCTCCCGGTTCAAGTGAATTCCTCTCCTCAG 119293

QY 1129 CCTCTGAGTAGCTGGGA 1146
|||||
```

```
DB 119294 CCTCCCGAGTAGCTGGGA 119311

RESULT 7
US-09-292-542A-1/c
; Sequence 1, Application US/09292542A
; Patent No. 6531279
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Genomic Sequence Of The 5-Lipoxygenase-Activating Protein (FLAP)
; Patent No. 6531279
; TITLE OF INVENTION: Polymorphic Markers Thereof And Methods For Detection Of Asthma.
; FILE REFERENCE: GENSET.026A
; CURRENT FILING DATE: 1999-04-15
; CURRENT APPLICATION NUMBER: US/09/292,542A
; PRIOR APPLICATION NUMBER: US 60/081893
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: US 60/091314
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/123406
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 43069
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..7708
; OTHER INFORMATION: potential 5' regulatory region
; NAME/KEY: misc_feature
; LOCATION: 36604..43069
; OTHER INFORMATION: potential 3' regulatory region
; NAME/KEY: exon
; LOCATION: 7709..7852
; OTHER INFORMATION: exon1
; NAME/KEY: exon
; LOCATION: 16236..16335
; OTHER INFORMATION: exon2
; NAME/KEY: exon
; LOCATION: 24227..24297
; OTHER INFORMATION: exon3
; NAME/KEY: exon
; LOCATION: 28133..28214
; OTHER INFORMATION: exon4
; NAME/KEY: exon
; LOCATION: 36128..36605
; OTHER INFORMATION: exon5
; NAME/KEY: misc_feature
; LOCATION: 7783..7785
; OTHER INFORMATION: ATG
; NAME/KEY: misc_feature
; LOCATION: 36288..36290
; OTHER INFORMATION: stop : TAA
; NAME/KEY: polyA signal
; LOCATION: 36581..36586
; OTHER INFORMATION: AATAAA
; NAME/KEY: misc_feature
; LOCATION: 7008..8116
; OTHER INFORMATION: homology with sequence in ref genbank : M60470
; NAME/KEY: misc_feature
; LOCATION: 15995..16549
; OTHER INFORMATION: homology with sequence in ref genbank : M63259
; NAME/KEY: misc_feature
; LOCATION: 24059..24597
; OTHER INFORMATION: homology with sequence in ref genbank : M63260
; NAME/KEY: misc_feature
; LOCATION: 27873..28412
; OTHER INFORMATION: homology with sequence in ref genbank : M63261
; NAME/KEY: misc_feature
```

```
LOCATION: 35977..36926
OTHER INFORMATION: homology with sequence in ref genbank : M63262
NAME/KEY: misc feature
LOCATION: 7613
OTHER INFORMATION: diverging nucleotide deletion of a A in ref : M60470
NAME/KEY: misc feature
LOCATION: 16347
OTHER INFORMATION: diverging nucleotide G in ref : M63259
NAME/KEY: misc feature
LOCATION: 16348
OTHER INFORMATION: diverging nucleotide A in ref : M63259
NAME/KEY: misc feature
LOCATION: 24060
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260
NAME/KEY: misc feature
LOCATION: 24067
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63260
NAME/KEY: misc feature
LOCATION: 27903
OTHER INFORMATION: diverging nucleotide deletion of a C in ref : M63261
NAME/KEY: misc feature
LOCATION: 28327
OTHER INFORMATION: diverging nucleotide deletion of a G in ref : M63261
NAME/KEY: misc feature
LOCATION: 3851..4189
OTHER INFORMATION: 10-517
NAME/KEY: misc feature
LOCATION: 4120..4390
OTHER INFORMATION: 10-518
NAME/KEY: misc feature
LOCATION: 4373..4792
OTHER INFORMATION: 10-253
NAME/KEY: misc feature
LOCATION: 4814..5043
OTHER INFORMATION: 10-499
NAME/KEY: misc feature
LOCATION: 4956..5422
OTHER INFORMATION: 10-500
NAME/KEY: misc feature
LOCATION: 5324..5996
OTHER INFORMATION: 10-522
NAME/KEY: misc feature
LOCATION: 6218..6672
OTHER INFORMATION: 10-503
NAME/KEY: misc feature
LOCATION: 6522..6790
OTHER INFORMATION: 10-504
NAME/KEY: misc feature
LOCATION: 7120..7574
OTHER INFORMATION: 10-204
NAME/KEY: misc feature
LOCATION: 7513..7933
OTHER INFORMATION: 10-32
NAME/KEY: misc feature
LOCATION: 16114..16533
OTHER INFORMATION: 10-33
NAME/KEY: misc feature
LOCATION: 24072..24425
OTHER INFORMATION: 10-34
NAME/KEY: misc feature
LOCATION: 27978..28401
OTHER INFORMATION: 10-35
NAME/KEY: misc feature
LOCATION: 36020..36465
OTHER INFORMATION: 10-36
NAME/KEY: misc feature
LOCATION: 36318..36669
OTHER INFORMATION: 10-498
NAME/KEY: misc feature
LOCATION: 38441..38840
OTHER INFORMATION: 12-629
NAME/KEY: misc feature
LOCATION: 42233..42749
```

```
OTHER INFORMATION: 12-628 complement
NAME/KEY: allele
LOCATION: 3950
OTHER INFORMATION: 10-517-100 : polymorphic base S
NAME/KEY: allele
LOCATION: 4243
OTHER INFORMATION: 10-518-125 : polymorphic base K
NAME/KEY: allele
LOCATION: 4312
OTHER INFORMATION: 10-518-194 : polymorphic base R
NAME/KEY: allele
LOCATION: 4490
OTHER INFORMATION: 10-253-118 : polymorphic base R
NAME/KEY: allele
LOCATION: 4670
OTHER INFORMATION: 10-253-298 : polymorphic base S
NAME/KEY: allele
LOCATION: 4687
OTHER INFORMATION: 10-253-315 : polymorphic base Y
NAME/KEY: allele
LOCATION: 4968
OTHER INFORMATION: 10-499-155 : polymorphic base R
NAME/KEY: allele
LOCATION: 5140
OTHER INFORMATION: 10-500-185 : polymorphic base Y
NAME/KEY: allele
LOCATION: 5213
OTHER INFORMATION: 10-500-258 : polymorphic base K
NAME/KEY: allele
LOCATION: 5364
OTHER INFORMATION: 10-500-410 : polymorphic base R
NAME/KEY: allele
LOCATION: 5594
OTHER INFORMATION: 10-522-71 : polymorphic base R
NAME/KEY: allele
LOCATION: 6370
OTHER INFORMATION: 10-503-159 : polymorphic base K
NAME/KEY: allele
LOCATION: 6693
OTHER INFORMATION: 10-504-172 : polymorphic base W
NAME/KEY: allele
LOCATION: 6763
OTHER INFORMATION: 10-504-243 : polymorphic base M
NAME/KEY: allele
LOCATION: 7445
OTHER INFORMATION: 10-204-326 : polymorphic base R
NAME/KEY: allele
LOCATION: 7870
OTHER INFORMATION: 10-32-357 : polymorphic base M
NAME/KEY: allele
LOCATION: 16288
OTHER INFORMATION: 10-33-175 : polymorphic base Y
NAME/KEY: allele
LOCATION: 16347
OTHER INFORMATION: 10-33-234 : polymorphic base M
NAME/KEY: allele
LOCATION: 16383
OTHER INFORMATION: 10-33-270 : polymorphic base R
NAME/KEY: allele
```

```
Query Match          9.6%; Score 110.4; DB 4; Length 43069;
Best Local Similarity 85.4%; Pred. No. 9.2e-24;
Matches 123; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1004 CTGAATCTTTTCTTTTATTTTTCAGATGGAGTCTGCTCTGTGGCCAGCTGGAGTG 1063
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34796 CTTTTCCTTTCTTTTCTTTTTCAGACTGAGTCAAGCTCTATTGCCACCTGGAGTG 34737

QY 1064 CAGTGGTGTGATCTCAGCTTACTGCAACTCTGTCTCCCGGGTTCAAGCAATTCTCCCAT 1123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34736 CAGTGGCAGATCTCAGATCACTGCAACTCTGTCTCCCGGGTTCAAGCGATTCTCTGC 34677

QY 1124 CTCAGCTCTCTAGTAGTGGGAT 1147
```

[illegible]

```

; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 10-520-256 : polymorphic base C or T
; NAME/KEY: misc binding
; LOCATION: 458..477
; OTHER INFORMATION: 10-520-256.mis1, potential
; NAME/KEY: primer bind
; LOCATION: 294..310
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 743..760
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 466..490
; OTHER INFORMATION: 10-520-256 potential probe
; US-09-641-638-5

Query Match          9.3%; Score 107; DB 4; Length 955;
Best Local Similarity 77.6%; Pred. No. 1.2e-23;
Matches 142; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

QY 965 TGGAGTTTGATTATCCCTCTCCCTCCCAACACTAGGAACCTGAAATCTTTTCTTTTATT 1024
Db 738 TGGAGCTAGAGAGCTTTCTCTCTCTCTCTCCGCCCTTCTCTCTTTTCTTTTCTTTT 679

QY 1025 TTTTGAGTGGAGTCTTGCTCTGTGTCCTCCCA-GCTGGAGTGCAGTGTGTGATCTCAGCTT 1083
Db 678 TTTTGAGCGGAGTCTTGCTCTGTGTCCTCCCAAGCTGGAGTGCAGTGTGCGGATCTCAGCTC 619

QY 1084 ACTGCAACTCTGTCTCCCGGGTTCAAGCAATTTCTCCCATCTCAGCTCTCCTGAGTAGCTG 1143
Db 618 ACTGCAAGCTCTGCTCCCGGGTTCAAGCAATTTCTCTCTCTCAGCTCTCTCAGTAGCTG 559

QY 1144 GGA 1146
Db 558 GGA 556

RESULT 11
US-09-641-638-6/c
; Sequence 6, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 6
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 10-500-258 : polymorphic base G or T
; NAME/KEY: misc binding
; LOCATION: 458..477
; OTHER INFORMATION: 10-500-258.mis1, potential
; NAME/KEY: primer bind
; LOCATION: 221..237
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 670..687
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 466..490
; OTHER INFORMATION: 10-500-258 potential probe
; US-09-641-638-6

Query Match          9.3%; Score 107; DB 4; Length 955;
Best Local Similarity 77.6%; Pred. No. 1.2e-23;
Matches 142; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

QY 965 TGGAGTTTGATTATCCCTCTCCCTCCCAACACTAGGAACCTGAAATCTTTTCTTTTATT 1024
Db 665 TGGAGCTAGAGAGCTTTCTCTCTCTCTCTCTCCGCCCTTCTCTCTTTTCTTTTCTTTT 606

QY 1025 TTTTGAGTGGAGTCTTGCTCTGTGTCCTCCCA-GCTGGAGTGCAGTGTGTGATCTCAGCTT 1083
Db 605 TTTTGAGCGGAGTCTTGCTCTGTGTCCTCCCAAGCTGGAGTGCAGTGTGCGGATCTCAGCTC 546

QY 1084 ACTGCAACTCTGTCTCCCGGGTTCAAGCAATTTCTCCCATCTCAGCTCTCCTGAGTAGCTG 1143
Db 545 ACTGCAAGCTCTGCTCCCGGGTTCAAGCAATTTCTCTCTCTCAGCTCTCTCAGTAGCTG 486

QY 1144 GGA 1146
Db 485 GGA 483

RESULT 12
US-09-641-638-7/c
; Sequence 7, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 7
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 10-500-410 : polymorphic base A or G
; NAME/KEY: misc binding
; LOCATION: 458..477
; OTHER INFORMATION: 10-500-410.mis1, potential
; NAME/KEY: primer bind
; LOCATION: 221..237
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 670..687
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 466..490
; OTHER INFORMATION: 10-500-258 potential probe
; US-09-641-638-6
```



```

; NAME/KEY: misc binding
; LOCATION: 479..498
; OTHER INFORMATION: 10-500-410.mis2, potential complement
; NAME/KEY: primer bind
; LOCATION: 70..86
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer bind
; LOCATION: 519..536
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc binding
; LOCATION: 466..490
; OTHER INFORMATION: 10-500-410 potential probe
US-09-641-638--7

Query Match          9.3%; Score 107; DB 4; Length 955;
Best Local Similarity 77.6%; Pred.No.1.2e-23;
Matches 142; Conservative 0; Mismatches 40; Indels 1; Gaps 1;

QY      965 TGGAGTTTGATTATCCCTCTCTCCCCAACACCTAGGAACCTGATCTTTCCTTTTATT 1024
DB       514 TGGAGCTAGAGAGCTTTCTCTCTCTCTCCTCCCGCYCTTCCTCTCTTTTTTTTTTT 455

QY      1025 TTTTGAGATCGAGCTCTGTCTCTGTGGCCCA-GCTGGAGTGCAGTGGTGTGATCTCAGCTT 1083
DB       454 TTTTGAACCGAGGCTTGCTCTGTGCCAGGCTGGAGTGCAGTGGCGCATCTCAGCTC 395

QY      1084 ACTGCAACCTCTGTCTCCCGGGTTCAAGCAATTCTCCCATTCTCAGCCTCCTGAGTAGCTG 1143
DB       394 ACTGCAAGCTCTGCCTCCCGGGTTCAGCCATTCTCTGTCTCAGCCTCCTCAGTAGCTG 335

QY      1144 GGA 1146
DB       334 GGA 332

RESULT 13
US-09-641-638-650/c
; Sequence 650, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilva
; APPLICANT: Cohen, Amick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 650
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-522-71 : polymorphic base A or G
; NAME/KEY: misc binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-522-71.misl, potential
; NAME/KEY: misc binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-522-71.mis2, potential complement
```

[illegible]

; ORGANISM: Homo sapiens  
US-09-489-847-22  
Query Match 9.2%; Score 106; DB 4; Length 2876;  
Best Local Similarity 85.5%; Pred. No. 4.6e-23;  
Matches 118; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1009 TCTTTTCTTTTATTTTGGAGATGGAGTCTTGTCTGTTGCCAGCTGGAGTGCAGTG 1068  
DB 2873 TTTTCTTTTATTTTGGAGATGGAGTCTTGTCTGTTGCCAGCTGGAGTGCAGTG 2814  
QY 1069 GTGTGATCTCAGCTTACTGCAACCTCTGTCTCCCGGTTCAAGCAATTCTCCCATCTCAG 1128  
DB 2813 GAGCAATCTCAGCTCAGTGCAGCTCCGCTCCAGGTTCAAGCAATTCTCTGCCTCAG 2754  
QY 1129 CCTCCTGAGTAGCTGGGA 1146  
DB 2753 CCTCCTGAGTAGCTGGGA 2736  
Search completed: December 1, 2003, 09:38:32  
Job time : 95 secs

; NAME/KEY: SITE  
; LOCATION: (1284)  
; OTHER INFORMATION: n equals a.t.g, or c  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1287)  
; OTHER INFORMATION: n equals a.t.g, or c  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1378)  
; OTHER INFORMATION: n equals a.t.g, or c  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1912)  
; OTHER INFORMATION: n equals a.t.g, or c  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1913)  
; OTHER INFORMATION: n equals a.t.g, or c  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1947)  
; OTHER INFORMATION: n equals a.t.g, or c  
US-09-489-847-111  
Query Match 9.2%; Score 106; DB 4; Length 2871;  
Best Local Similarity 85.5%; Pred. No. 4.6e-23;  
Matches 118; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1009 TCTTTTCTTTTATTTTGGAGATGGAGTCTTGTCTGTTGCCAGCTGGAGTGCAGTG 1068  
DB 2868 TTTTCTTTTATTTTGGAGATGGAGTCTTGTCTGTTGCCAGCTGGAGTGCAGTG 2809  
QY 1069 GTGTGATCTCAGCTTACTGCAACCTCTGTCTCCCGGTTCAAGCAATTCTCCCATCTCAG 1128  
DB 2808 GAGCAATCTCAGCTCAGTGCAGCTCCGCTCCAGGTTCAAGCAATTCTCTGCCTCAG 2749  
QY 1129 CCTCCTGAGTAGCTGGGA 1146  
DB 2748 CCTCCTGAGTAGCTGGGA 2731

RESULT 15  
US-09-489-847-22/c  
; Sequence 22, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031PI  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 2876  
; TYPE: DNA

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2003, 07:30:33 ; Search time 1865 Seconds  
(without alignments)  
2026.279 Million cell updates/sec

Title: US-10-049-742-22

Perfect score: 1147  
Sequence: 1 gcttgggtcaagcagaata.....gctctctgagtagctgggat 1147

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1547345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09D\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	1146	99.9	3090	12	US-09-814-353-19835
C 2	1125	98.1	2945	14	US-10-198-846-12415
C 3	743	64.8	754	12	US-10-096-534-43
C 4	743	64.8	754	14	US-10-097-340-182
C 5	453	39.5	7453	11	US-09-764-891-8949
C 6	293.2	25.6	443	12	US-09-814-353-14171
C 7	257	22.4	259	12	US-09-814-353-1424
C 8	257	22.4	259	12	US-09-814-353-7786
C 9	251.4	21.9	12572	14	US-10-228-406A-10
C 10	251.4	21.9	14078	12	US-10-134-288-1
C 11	251.4	21.9	16713	14	US-10-228-406A-9
C 12	234.6	20.5	339	10	US-09-867-701-8500
C 13	137.6	12.0	510	12	US-09-814-353-4106
C 14	137.6	12.0	510	12	US-09-814-353-10412
C 15	137	11.9	598	12	US-09-814-353-16796
C 16	137	11.9	621	14	US-10-198-846-9074

C 17	116.8	10.2	28001	13	US-10-193-295-3	Sequence 3, Appli
C 18	115	10.0	403	12	US-10-027-632-279852	Sequence 279852,
C 19	115	10.0	403	12	US-10-027-632-279852	Sequence 279852,
C 20	114	9.9	48763	14	US-10-282-048-3	Sequence 3, Appli
C 21	113.6	9.9	498	10	US-09-520-300A-1487	Sequence 1487, Ap
C 22	113.6	9.9	498	12	US-10-099-926-1487	Sequence 1487, Ap
C 23	113.6	9.9	498	13	US-10-033-528-1487	Sequence 1487, Ap
C 24	112.8	9.8	2147	10	US-09-764-847-1738	Sequence 1738, Ap
C 25	112.8	9.8	2147	14	US-10-092-154-1738	Sequence 1738, Ap
C 26	112.8	9.8	32174	9	US-09-988-711-158	Sequence 158, App
C 27	112.8	9.8	32174	9	US-09-764-860-1134	Sequence 1134, Ap
C 28	112.8	9.8	32174	10	US-09-764-877-2645	Sequence 2645, Ap
C 29	112.8	9.8	32174	10	US-09-560-670-232	Sequence 232, App
C 30	112.8	9.8	32174	10	US-09-764-904-90	Sequence 90, Appli
C 31	112.8	9.8	32174	11	US-09-764-891-6490	Sequence 6480 Ap
C 32	112.8	9.8	32174	11	US-09-764-891-10135	Sequence 10135, A
C 33	112.8	9.8	32174	11	US-09-764-891-10179	Sequence 10179, A
C 34	112.8	9.8	32174	12	US-10-212-872-1134	Sequence 1134, Ap
C 35	112.8	9.8	32174	14	US-10-091-548-90	Sequence 90, Appli
C 36	112.8	9.8	32174	14	US-10-074-095-1134	Sequence 1134, Ap
C 37	112.4	9.8	174493	10	US-09-804-471A-3	Sequence 3, Appli
C 38	112.4	9.8	174493	14	US-10-238-709-3	Sequence 3, Appli
C 39	112.2	9.8	12149	9	US-09-764-869-2258	Sequence 2258, Ap
C 40	112.2	9.8	12149	14	US-10-091-504-2258	Sequence 2258, Ap
C 41	112.2	9.8	106344	11	US-09-910-185-10	Sequence 10, Appli
C 42	112	9.8	689	12	US-10-027-632-135995	Sequence 135995,
C 43	112	9.8	689	13	US-10-027-632-135995	Sequence 135995,
C 44	112	9.8	1536	12	US-10-027-632-252944	Sequence 252944,
C 45	112	9.8	1536	13	US-10-027-632-252944	Sequence 252944,

ALIGNMENTS

RESULT 1

US-09-814-353-19835/c  
; Sequence 19835, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 22037  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 19835  
; LENGTH: 3090  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 488, 525  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-19835

Query Match

99.9%; Score 1146; DB 12; Length 3090;

Best Local Similarity 99.9%; Pred. No. 0; Matches 1146; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	GCCTTGGGTCAAGCAGAGATATTAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGGCTA	60						
Db	1667	GCCTTGGGTCAAGCAGAGATATTAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGGCTA	1608						
Qy	61	CTGCCAGCCTCAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCTCTGAGGATGA	120						
Db	1607	CTGCCAGCCTCAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCTCTGAGGATGA	1548						
Qy	121	GCTAAACCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTTGAATCAAGAA	180						
Db	1547	GCTAAACCTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTTGAATCAAGAA	1488						
Qy	181	GCCTATAGACAGCTGGGAGTATGTTTCCATCCTGACAAAATCATCATCCCGGGCTGA	240						
Db	1487	GCCTATAGACAGCTGGGAGTATGTTTCCATCCTGACAAAATCATCATCCCGGGCTGA	1428						
Qy	241	GGAGGCTTCAAGGTTTTCGGAGCAGCTTGGACATTTGTCAAGATGCTGAAAACGAAA	300						
Db	1427	GGAGGCTTCAAGGTTTTCGGAGCAGCTTGGACATTTGTCAAGATGCTGAAAACGAAA	1368						
Qy	301	GGAGTATGAGATGAACGAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTCT	360						
Db	1367	GGAGTATGAGATGAACGAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTCT	1308						
Qy	361	GTCCAAAGCTGCAAGATGACCTCAAGAGGCAATGATATATGATGTAGCCGATGCCA	420						
Db	1307	GTCCAAAGCTGCAAGATGACCTCAAGAGGCAATGATATATGATGTAGCCGATGCCA	1248						
Qy	421	AGGAAGCATAGAGGTTTGAATGACCGGAACTTAAGTGGCCAGATCTGTGCTGA	480						
Db	1247	AGGAAGCATAGAGGTTTGAATGACCGGAACTTAAGTGGCCAGATCTGTGCTGA	1188						
Qy	481	GTGTAATAGCTGCAATCTCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTTGG	540						
Db	1187	GTGTAATAGCTGCAATCTCTGCTGAGGAAGGAGACTTTTGGCAGAGTCAAGCATGTTGG	1128						
Qy	541	CCTCAAGATCACTTACTTTTGCACTGATGATGGAAGGTTGATGACATCAGAGATGGGC	600						
Db	1127	CCTCAAGATCACTTACTTTTGCACTGATGATGGAAGGTTGATGACATCAGAGATGGGC	1068						
Qy	601	TGATGCCAGCTGTAGGATATCTCCACAGATCCACAGAGTCCCTATCAGATCTCAT	660						
Db	1067	TGATGCCAGCTGTAGGATATCTCCACAGATCCACAGAGTCCCTATCAGATCTCAT	1008						
Qy	661	TGATCTCGGATTCAGGACACAGAGGGCGGAGAGAGCCACCCAGATGCCCTCTCTGC	720						
Db	1007	TGATCTCGGATTCAGGACACAGAGGGCGGAGAGAGCCACCCAGATGCCCTCTCTGC	948						
Qy	721	TGATCTCGGATTCAGGATCGGATCTTTCAAGTACCCGAGGCGAGATGCCCAATGG	780						
Db	947	TGATCTCGGATTCAGGATCGGATCTTTCAAGTACCCGAGGCGAGATGCCCAATGG	888						
Qy	781	GAACTTTCTTGAGCTCTCAGCTGCCCTCGAGCGGCTGAGCTCTTAAGCCCAACAG	840						
Db	887	GAACTTTCTTGAGCTCTCAGCTGCCCTCGAGCGGCTGAGCTCTTAAGCCCAACAG	828						
Qy	841	CACAGTACCAAGGAGAGCCAAAACCTAAGCGGGCGAAGAAAGTGAAGAGGCCCTTCCA	900						
Db	827	CACAGTACCAAGGAGAGCCAAAACCTAAGCGGGCGAAGAAAGTGAAGAGGCCCTTCCA	768						
Qy	901	ACGTTGATGCCCTTCTCTTTCTCAAAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAC	960						
Db	767	ACGTTGATGCCCTTCTCTTTCTCAAAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAC	708						
Qy	961	AGGATGGAGTTGATTTATCCCTCTCCCAACACCTAGGAACTGAATCTTTTCTTTT	1020						
Db	707	AGGATGGAGTTGATTTATCCCTCTCCCAACACCTAGGAACTGAATCTTTTCTTTT	648						
Qy	1021	TATTTTTGAGATGGAGTCTTGCTGTGTCGCCAGCTGGAGTGACGTGGTGTGATCTCAG	1080						

Db	647	TATTTTTGAGATGGAGTCTTGCTGTGTCGCCAGCTGGAGTGCGATGCTCTCAG	588						
Qy	1081	CTTACTGCAACCTCTGTCTCCCGGTTCAAGCAATTTCCCATCTTACGCTCTCTGAGTAG	1140						
Db	587	CTTACTGCAACCTCTGTCTCCCGGTTCAAGCAATTTCCCATCTTACGCTCTCTGAGTAG	528						
Qy	1141	CTGGAT 1147							
Db	527	CTNGAT 521							

RESULT 2  
US-10-198-846-12415/c  
; Sequence 12415, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10198.846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12415  
; LENGTH: 2945  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-198-846-12415

Query Match 98.1%; Score 1125; DB 14; Length 2945;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1147; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy	1	GCCTTGGGTCAAGCAGAGATATTAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGGCTA	60						
Db	1673	GCCTTGGGTCAAGCAGAGATATTAATAGCAGGGGAATGCACCTGTAGCTAGTGGCGGCTA	1614						
Qy	61	CTGCCAGCCTCAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCTCTGAGGATGA	120						
Db	1613	CTGCCAGCCTCAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGGTTCTCTGAGGATGA	1554						
Qy	121	GCTAAACCTTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTTGAATCAAGAA	180						
Db	1553	GCTAAACCTTTTCCATGTACTGGGGTTGAGGCCACAGCATCAGATGTTTGAATCAAGAA	1494						
Qy	181	GGCCTATAGACAGCTGGCAGTGAATGTTTCATCTGACAAAAATCATCATCCCGGGCTGA	240						
Db	1493	GGCCTATAGACAGCTGGCAGTGAATGTTTCATCTGACAAAAATCATCATCCCGGGCTGA	1434						
Qy	241	GGAGCCTTCAAGGTTTTCGGAGCAGCTTGGACATTTGTCAAGATGCTGAAAACGAAA	300						
Db	1433	GGAGCCTTCAAGGTTTTCGGAGCAGCTTGGACATTTGTCAAGATGCTGAAAACGAAA	1374						
Qy	301	GGAGTATGAGATGAACGAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTTCT	360						
Db	1373	GGAGTATGAGATGAACGAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTTCT	1314						
Qy	361	GTCCAAAGCTGCAAGATGACCTCAAGAGGCAATGAATATCTATGATGTAGCCGATGCCA	420						
Db	1313	GTCCAAAGCTGCAAGATGACCTCAAGAGGCAATGAATATCTATGATGTAGCCGATGCCA	1254						
Qy	421	AGGAAAGCATAGGAGTTTGAATGGACGGGAACTTAAGAGTGGCAGATCTGTGCTGA	480						
Db	1253	AGGAAAGCATAGGAGTTTGAATGGACGGGAACTTAAGAGTGGCAGATCTGTGCTGA	1194						
Qy	481	GTGTAATAGCTGCAATCTCTGCTGAGGAAGGAGACTTTTGGGCGAGAGTCAAGCATGTTGG	540						

Db 1193 GTGTAATAGGCTGATCTCTGCTAGGAAGAGATCTTTGGCCAGATCAAGCATGTTGGG 1134  
Qy 541 CCTCAAGATCACCTACTTTGACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGG 600  
Db 1133 CCTCAAGATCACCTACTTTGACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGG 1074  
Qy 601 TGGATGCCAGCGTGTAGTATCTCCAGATACCCAGAGTCCCTATCAATCTCATTT 660  
Db 1073 TGGATGCCAGCGTGTAGTATCTCCAGATACCCAGAGTCCCTATCAATCTCATTT 1014  
Qy 661 TGGTCTCGAATTCAGGACACAGAGGGCGGAGAGGCCACCCAGATGCCCTCTCTGC 720  
Db 1013 TGGTCTCGAATTCAGGACACAGAGGGCGGAGAGGCCACCCAGATGCCCTCTCTGC 954  
Qy 721 TGATCTTCAGATTTCTTGAGTGGATCTTTCAAGTACCCCGGAGAGTCCCAATGG 780  
Db 953 TGATCTTCAGATTTCTTGAGTGGATCTTTCAAGTACCCCGGAGAGTCCCAATGG 894  
Qy 781 GAACTTCTTTCAGCTCTCAGCTGCGCTGCGCTGGAGCGCTGCAGCTCTAAGGCCAACAG 840  
Db 893 GAACTTCTTTCAGCTCTCAGCTGCGCTGCGCTGGAGCGCTGCAGCTCTAAGGCCAACAG 834  
Qy 841 CACAGTACCCAGGAGAGCAAACTTAAGCGGGGAGAAAGTGAAGGCGCTTCCA 900  
Db 833 CACAGTACCCAGGAGAGCAAACTTAAGCGGGGAGAAAGTGAAGGCGCTTCCA 774  
Qy 901 ACGTGTAGTCCCTCTCTCTTCTCAATCAATGTCAGGAGTCAAAAGGCTGTAGCAC 960  
Db 773 ACGTGTAGTCCCTCTCTCTTCTCAATCAATGTCAGGAGTCAAAAGGCTGTAGCAC 714  
Qy 961 AGGATGGAGTTTGATTTATCCCTCTCCCGGAGAGTCAAGTGAATCTTTTCTTTT 1020  
Db 713 AGGATGGAGTTTGATTTATCCCTCTCCCGGAGAGTCAAGTGAATCTTTTCTTTT 654  
Qy 1021 TATTTTTAGATGAGTCTTGCTCT-GTGCCCCAGCTGGAGTGAAGTGAATCTCA 1079  
Db 653 TATTTTTAGATGAGTCTTGCTCTGTTGTCAGCTGGAGTGAAGTGAATCTCA 594  
Qy 1080 GCTTACTGCACTCTGCTCCCGG-TTCAAGCAATTTCTCCATCTCAGCTCTCTGAGT 1138  
Db 593 GCTTACTGCACTCTGCTCCCGGCTTCAAGCAATTTCTCCATCTCAGCTCTCTGAGT 534  
Qy 1139 AGCTGGAT 1147  
Db 533 AGCTGGAT 525

RESULT 3  
US-10-096-534-43  
; Sequence 43, Application US/10096534  
; Publication No. US20030166887A1  
; GENERAL INFORMATION:  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Yates, Karen  
; APPLICANT: Mizuno, Shuichi  
; APPLICANT: Glowacki, Julie  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS  
; FILE REFERENCE: B0801/7244/KA/ERP  
; CURRENT APPLICATION NUMBER: US/10/096,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,980  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 43  
; LENGTH: 754  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-096-534-43  
Query Match 64.8%; Score 743; DB 12; Length 754;  
Best Local Similarity 99.9%; Pred. No. 1.7e-232;

Matches 754; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 301 GGAGTATGAGATGAAACGAAATGGCAGAGAATAGCTGAGCGGTGAGTAAATGATTTCT 360  
Db 1 GGAGTATGAGATGAAACGAAATGGCAGAGAATAGCTGAGCGGTGAGTAAATGATTTCT 60  
Qy 361 GTCCAACTGCAAGATCAACCTCAAGGAGGCAATGAATACTATGATGTAGCCGATGCCA 420  
Db 61 GTCCAACTGCAAGATCAACCTCAAGGAGGCAATGAATACTATGATGTAGCCGATGCCA 120  
Qy 421 AGGAAAGCATAGGAGTTTGAATGGACCGGAACTTAAAGTGGCAGATCAAGCATGTTGGG 480  
Db 121 AGGAAAGCATAGGAGTTTGAATGGACCGGAACTTAAAGTGGCAGATCAAGCATGTTGGG 180  
Qy 481 GTGTAATAGGCTGATCTGCTGAGGAAGAGACTTTTGGCAGAGTCAAGCATGTTGGG 540  
Db 181 GTGTAATAGGCTGATCTGCTGAGGAAGAGACTTTTGGCAGAGTCAAGCATGTTGGG 240  
Qy 541 CCTCAAGATCACCTTACCTTTCACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGG 600  
Db 241 CCTCAAGATCACCTTACCTTTCACCTGATGGATGGAAGGTGTATGACATCAAGAGTGGG 300  
Qy 601 TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTTATCATCTCATTT 660  
Db 301 TGGATGCCAGCGTGTAGTATCTCCCGAGATACCCAGAGTCCCTTATCATCTCATTT 360  
Qy 661 TGGTCTCGGATTCAGGACCAAGGCGGCGAGAGCCACCCAGATGCCCTCTCTGC 720  
Db 361 TGGTCTCGGATTCAGGACCAAGGCGGCGAGAGCCACCCAGATGCCCTCTCTGC 420  
Qy 721 TGATCTTCAGATTTCTTGAGTGGATCTTTCAAGTACCCCGGAGAGTCCCAATGG 780  
Db 421 TGATCTTCAGATTTCTTGAGTGGATCTTTCAAGTACCCCGGAGAGTCCCAATGG 479  
Qy 781 GAACTTCTTTCAGCTCTCTGAGCTGCCCTGGAGCGCTGAGCGCTCTAAGCCCAACAG 840  
Db 480 GAACTTCTTTCAGCTCTCTGAGCTGCCCTGGAGCGCTGAGCGCTCTAAGCCCAACAG 539  
Qy 841 CACAGTACCCAGGAGAGCAAACTTAAGCGGGGAGAAAGTGAAGGCGCTTCCA 900  
Db 540 CACAGTACCCAGGAGAGCAAACTTAAGCGGGGAGAAAGTGAAGGCGCTTCCA 599  
Qy 901 ACGTGTAGTCCCTCTCTCTTTCTCAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAC 960  
Db 600 ACGTGTAGTCCCTCTCTTTCTCTCAATCAATGTCAAGGAGTCAAAAGGCTGTAGCAC 659  
Qy 961 AGGATGGAGTTTGATTTATCCCTCTCCCGCAACCTAGGAGTGAATCTTTTCTTTT 1020  
Db 660 AGGATGGAGTTTGATTTATCCCTCTCCCGCAACCTAGGAGTGAATCTTTTCTTTT 719  
Qy 1021 TATTTTTAGATGAGTCTTGCTCTGTTGCCAG 1055  
Db 720 TATTTTTAGATGAGTCTTGCTCTGTTGCCAG 754

RESULT 4  
US-10-097-340-182  
; Sequence 182, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU

```

; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GUATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 182
; LENGTH: 754
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-097-340-182

```

Query Match	64.8%	Score 743	DB 14	Length 754
Best Local Similarity	99.9%	Pred. No. 1.7e-232		
Matches 754	Conservative 0	Mismatches 0	Indels 1	Gaps 1
QY	301	GGAGTATGAGATGAAACGAATGCGCAGAGATGAGCTGAGCGGTCACTAAATGAGTTTCT	360	
DB	1	GGAGTATGAGATGAAACGAATGCGCAGAGATGAGCTGAGCGGTCACTAAATGAGTTTCT	60	
QY	361	GTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGAATATCTATGTGTAGCGGATGCCA	420	
DB	61	GTCCAAAGCTGCAAGATGACCTCAAGGAGGCAATGAATATCTATGTGTAGCGGATGCCA	120	
QY	421	AGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGTGCACAGATACTGTGCTGA	480	
DB	121	AGGAAGCATAGGAGGTTTGAATGGACCGGGAACCTTAAGAGTGCACAGATACTGTGCTGA	180	
QY	481	GTGTAATAGGCTGCACTCTGCTGAGGAGGAGACCTTTGGCAGAGTCAAGCATGTTGG	540	
DB	181	GTGTAATAGGCTGCACTCTGCTGAGGAGGAGACCTTTGGCAGAGTCAAGCATGTTGG	240	
QY	541	CCTCAAGATCACCTACTTTGGACTGTAGATGGAAGGTGTATGACATCACAGAGTGGC	600	
DB	241	CCTCAAGATCACCTACTTTGGACTGTAGATGGAAGGTGTATGACATCACAGAGTGGC	300	
QY	601	TGGATGCCAGCGTGTAGTATCTCTCCACAGATACCACACAGAGTCCCCCTATCACATCTCAT	660	
DB	301	TGGATGCCAGCGTGTAGTATCTCTCCACAGATACCACACAGAGTCCCCCTATCACATCTCAT	360	
QY	661	TGGTTCTCGGATCCAGGSCACAGAGGCGGCAGAGAGCCACCCACAGATGCCCTCCTCG	720	
DB	361	TGGTTCTCGGATCCAGGSCACAGAGGCGGCAGAGAGCCACCCACAGATGCCCTCCTCG	420	
QY	721	TGATCTTCAGGATTTCTTGAGTCCGATCTTTCAAGTACCCCAAGGCGAGATGCCCAATGG	780	
DB	421	TGATCTTCAGGATTTCTTGAGTCCGATCTTTCAAGTACCCCAAGGCGAGATGCCCAATGG	479	
QY	781	GAACTTCTTTGAGCTCCTCAGCTTGCCCTTGAGGCGGCTGCAGCCTCTAAGGCCCAACAG	840	
DB	480	GAACTTCTTTGAGCTCCTCAGCTTGCCCTTGAGGCGGCTGCAGCCTCTAAGGCCCAACAG	539	
QY	841	CACAGTACCACAGGAGAGGACCAACTTAAGCGGGGGAAGAAAGTGAAGGAGGCCCTTCCA	900	
DB	540	CACAGTACCACAGGAGAGGACCAACTTAAGCGGGGGAAGAAAGTGAAGGAGGCCCTTCCA	599	

```

901  ACGTTGATGCCCTTCTCTTTCTCAATGTCAGGAGTCAAAAGGGCTGTAGCAC 360
      |||||
600  ACGTTGATGCCCTTCTCTTTCTCAATGTCAGGAGTCAAAAGGGCTGTAGCAC 659
      |||||
961  AGATGGAGGTTTGATTTATCCCTCTCTCCCAACACCTAGGAACCTGAATCTTTTCTTTT 1020
      |||||
660  AGGATGGAGTTTGATTTATCCCTCTCTCCCAACACCTAGGAACCTGAATCTTTTCTTTT 719
      |||||
1021 TATTTTTCAGATGGAGTCTTGCTCTGTTGGCCAG 1055
      |||||
720 TATTTTTCAGATGGAGTCTTGCTCTGTTGGCCAG 754
      |||||

RESULT 5
US-09-764-891-8949
; Sequence 8949, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8949
; LENGTH: 7453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (350)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-764-891-8949

Query March 19 5% Score 453 DB 11 Length 7453;

```

	Query Match	39.5%;	Score 453;	DB 11;	Length 7453;
	Best Local Similarity	100.0%;	Prod. No. 6e-137;		
	Matches 453;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	695	AGAGCCACCCAGATGCCCTCTCTGTGATCTTCAGGATTTCTTTAGTCGGATCTTTCAA	754		
Db	3315	AGAGCCACCCAGATGCCCTCTCTGTGATCTTCAGGATTTCTTTAGTCGGATCTTTCAA	3374		
Qy	755	GTACCCCCAGGGCAGATGCCCAATGGGAACCTCTTTGCAGCTCTCTCAGCCTGCCCTTGGG	814		
Db	3375	GTACCCCCAGGGCAGATGCCCAATGGGAACCTCTTTGCAGCTCTCTCAGCCTGCCCTTGGG	3434		
Qy	815	GCGCCTGCAGCCTCTTAAGCCCAACAGCACAGTACCCAAAGGAGAACCCAAACCTTAAGCGG	874		
Db	3435	GCGCCTGCAGCCTCTTAAGCCCAACAGCACAGTACCCAAAGGAGAACCCAAACCTTAAGCGG	3494		
Qy	875	CGGAAGAAAGTGAGGAGGCCCTTCCAAAGTTGATGCCCTTCTCTTTTCCTCAAAATCAATG	934		
Db	3495	CGGAAGAAAGTGAGGAGGCCCTTCCAAAGTTGATGCCCTTCTCTTTTCCTCAAAATCAATG	3554		
Qy	935	TCAGGAGTCAAAAGGGCTGTAGCACAGGATGAGTTTGATTTATCCCTCTCTCCCCCAAC	994		
Db	3555	TCAGGAGTCAAAAGGGCTGTAGCACAGGATGAGTTTGATTTATCCCTCTCTCCCCCAAC	3614		
Qy	995	ACCTAGAACTGAATCTTTTCTTTTATTTTATTTTGTAGATGGAGTCTTGCTGTGGCCA	1054		
Db	3615	ACCTAGAACTGAATCTTTTCTTTTATTTTATTTTGTAGATGGAGTCTTGCTGTGGCCA	3674		
Qy	1055	GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTGTGCTCCGGGTTCAAGCAA	1114		
Db	3675	GCTGGAGTGCAGTGGTGTGATCTCAGCTTACTGCAACCTGTGCTCCGGGTTCAAGCAA	3734		
Qy	1115	TTCTCCCATCTCAGCCTCCTGAGTAGCTGGAT	1147		
Db	3735	TTCTCCCATCTCAGCCTCCTGAGTAGCTGGAT	3767		

```

RESULT 6
US-09-814-353-14171/c
; Sequence 14171, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-0068
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14171
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-14171

Query Match      25.6%; Score 293.2; DB 12; Length 443;
Best Local Similarity 98.7%; Pred. No. 3.2e-85;
Matches 306; Conservative 0; Mismatches 3; Indels 1; Gaps 0

QY 133 CCATGTACTGGGGTGTGAGCCACACATCAGA-TGTTGAACTGAAGAGGCGCTATAG
DB 426 CAAGGTACTGGGGTTAAGGCCACAGCATCAGATTCTTGAACTGAAGAGGCGCTATAG
QY 192 AGCTGGCAGTGATGGTTTCATCTTGACAAAATCATCATCCCGGCTGAGGAGGCGCT
DB 366 AGCTGGCAGTGATGGTTTCATCTTGACAAAATCATCATCCCGGCTGAGGAGGCGCT
QY 252 AGTTTTTCGCGAGCGCTTGGGACATTGTGAGCAATGCTGAAAGCGAAAGGAGTATGA
DB 306 AGTTTTTCGCGAGCGCTTGGGACATTGTGAGCAATGCTGAAAGCGAAAGGAGTATGA
QY 312 TGAACCAATGGCAGAGAATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTCCAAGC
DB 246 TGAACCAATGGCAGAGAATGAGCTGAGCGGTCAGTAAATGAGTTTCTGTCCAAGC
QY 372 AAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAAGGAAGAC
DB 186 AAGATGACCTCAAGGAGGCAATGAATACTATGATGTGTAGCCGATGCCAAGGAAGAC
QY 432 GGAGGTTTGA 441
DB 126 GGAGGTTTGA 117

RESULT 7
US-09-814-353-1424/c
; Sequence 1424, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

```

```

QY 464 GCACGATCTGTGCTGAGTGTAAATAGGCTGCATCTCTGTAGGAAGGAGACTTTTGGGCA 523
Db 504B GCACAUAUCUGUGAGUGUAUAAGGUGUGCAUCCUGUGAGGAAGGUGACUUUUGGGCA 5107
QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACTACTTTTGCACTGTAGTGGAAAGGTGTAT 583
Db 5108 GAGUCGAGCAUGTUGGGCCUAAAAUACCAUACUUTUGGCGUGAUGGAUGGAAGUGUAU 5167
QY 584 GACATCACAGATGGGCTGGATGCCAGCGGTAGTGTATCTCCACAGATACCCACAGAGTC 643
Db 5168 GAUAUCACAGAGUGGCGUGGAUGGCCAGGUGGUGGAUUCUCCCCAGAUACCCACAGAGUC 5227
QY 644 CCTATCACATCTCAATTTGGTTCTCGGATTCGAGGCA 680
Db 5228 CCUUGUCACAUCAUUAUUGUUAUACGGAUCCUUUCA 5264

RESULT 10
US-10-134-288-1
; Sequence 1, Application US/10134288
; Publication No. US20030165520A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Sheppard, Mike
; TITLE OF INVENTION: ATTENUATED FORMS OF BOVINE VIRAL DIARRHEA VIRUS
; FILE REFERENCE: FCI0435A
; CURRENT APPLICATION NUMBER: US/10/134,288
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/702,330
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/433,262
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 14078
; TYPE: DNA
; ORGANISM: Bovine Viral Diarrhea Virus
US-10-134-288-1

Query Match 21.9%; Score 251.4; DB 12; Length 14078;
Best Local Similarity 94.2%; Pred. No. 9.8e-71;
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 404 ATGTGTAGCCGATGCCAAGGAAAGCATAGAGAGTTTGAATGGACCGGAACTTAAGAGT 463
Db 4532 ATGTGCAGCCGATGCCAGGAAAGCATAGAGAGTTTGAATGGACCGGAACTTAAGAGT 4591
QY 464 GCCAGATACTGTGCTGAGTGTAAATAGGCTGCATCTCTGTAGGAAGGAGACTTTTGGGCA 523
Db 4592 GCCAGATACTGTGCTGAGTGTAAATAGGCTGCATCTCTGTAGGAAGGAGACTTTTGGGCA 4651
QY 524 GAGTCAAGCATGTTGGGCTCAAGATCACTACTTTTGCACTGTAGTGGAAAGGTGTAT 583
Db 4652 GAGTCGAGCATGTTGGGCTCAAAATCACCTACTTTGCGCTGTAGTGGAAAGGTGTAT 4711
QY 584 GACATCACAGATGGGCTGGATGCCAGCGGTAGTGTATCTCCACAGATACCCACAGAGTC 643
Db 4712 GATATCACAGATGGGCTGGATGCCAGCGTGTGGGAATCTCCACAGATACCCACAGAGTC 4771
QY 644 CCCTATCACATCTCAATTTGGTTCTCGGATTCGAGGCA 680
Db 4772 CCTTGTGCATCTCAATTTGGTTTCAGCGATGCCTTTCA 4808

RESULT 11
US-10-228-406A-9
; Sequence 9, Application US/10228406A
; Publication No. US20030104612A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Zylbarth, Gabriele

```

```

; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7786
; LENGTH: 259
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2
; OTHER INFORMATION: n = A,T,C or G
;
US-09-814-353-7786

Query Match 22.4%; Score 257; DB 12; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.7e-73;
Matches 257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 CTATAGACAGCTGGCAGTGTGTTTCATCTGTGACAAAATCATCATCCCGGGCTGAGGA 243
|||
Db 259 CTATAGACAGCTGGCAGTGTGTTTCATCTGTGACAAAATCATCATCCCGGGCTGAGGA 200
|||

Qy 244 GGCCTTCAAGGTTTTGGGAGCAGCTTGGGACATTGTGCAGCAATGCTGAAAAGCGAAAGGA 303
|||
Db 199 GGCCTTCAAGGTTTTGGGAGCAGCTTGGGACATTGTGCAGCAATGCTGAAAAGCGAAAGGA 140
|||

Qy 304 GTATGAGATGAAACGAAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTTCTGTC 363
|||
Db 139 GTATGAGATGAAACGAAATGGCAGAGAAATGAGCTGAGCCGGTCAGTAAATGAGTTTCTGTC 80
|||

Qy 364 CAAGCTGCAAGATGACCTCAAGGAGGCAATGAAATCTATCATGTGTAGCCGATGCCAAGG 423
|||
Db 79 CAAGCTGCAAGATGACCTCAAGGAGGCAATGAAATCTATCATGTGTAGCCGATGCCAAGG 20
|||

Qy 424 AAAGCATAGGAGTTTG 440
|||
Db 19 AAAGCATAGGAGTTTG 3
|||

RESULT 9
US-10-228-406A-10
; Sequence 10, Application US/10228406A
; Publication No. US20030104612A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Xuemei
; APPLICANT: Zybarch, Gabriele
; TITLE OF INVENTION: GENERATION OF TYPE I/TYPE II HYBRID FORM OF BOVINE
; TITLE OF INVENTION: VIRAL DIARRHEA VIRUS FOR USE AS VACCINE
; FILE REFERENCE: PCL1051A
; CURRENT APPLICATION NUMBER: US/10/228.406A
; CURRENT FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 12572
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid BVD virus NADL890
;
US-10-228-406A-10

Query Match 21.9%; Score 251.4; DB 14; Length 12572;
Best Local Similarity 72.2%; Pred. No. 9.2e-71;
Matches 200; Conservative 61; Mismatches 16; Indels 0; Gaps 0;

Qy 404 ATGTGTAGCCGATGCCAAGGAAAGCATAGGAGGTTTGAATGGACCGGAAACCTTAAGGT 463
|||
Db 498A AUGUGCAGCGGAUGCCAGGAAAGCAUAGGAGGJUUGAAAUAGGACCGGAAACCUAAGAGU 5047
|||

```



```

; TITLE OF INVENTION: GENERATION OF TYPE I/TYPE II HYBRID FORM OF BOVINE
; TITLE OF INVENTION: VIRAL DIARRHEA VIRUS FOR USE AS VACCINE
; FILE REFERENCE: PC11051A
; CURRENT APPLICATION NUMBER: US/10/228,406A
; CURRENT FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 16713
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pNADL890 vector
US-10-228-406A-9

Query Match          21.9%; Score 251.4; DB 14; Length 16713;
Best Local Similarity 94.2%; Pred. No. 1.1e-70;
Matches 261; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY      404  ATGTGTAGCCGATGCCAAGGAAGCATAGGAGGTTTGAATCGACCGGGAACTTAAGACT 463
Db      4988  ATGTGCAGCCGATGCCAGGGAAGCATAGGAGTTTGAATCGACCGGGAACTTAAGACT 5047

QY      464  GCCAGATACTGTGCTCGAGTGTAAATAGGCTGCATCTCTGCTGAGGAAGGACACATTTGGGCA 523
Db      5048  GCCAGATACTGTGCTCGAGTGTAAATAGGCTGCATCTCTGCTGAGGAAGTGAATTTGGGCA 5107

QY      524  GAGTCAAGCATGTTGGGCGCTCAGATCACTACTTTTGCCTGATGATCGAAAGGTGTAT 583
Db      5108  GAGTCCAGCATGTTGGGCGCTCAAAATCACTACTTTTGCCTGATGATCGAAAGGTGTAT 5167

QY      584  GACATCACAGATGGGCTGGATGCCACGGTGTAGGTATCTCCACAGATACCCACAGATC 643
Db      5168  GATATCACAGATGGGCTGGATGCCACGGTGTGGGAATCTCCACAGATACCCACAGATC 6227

QY      644  CCTATCACATCTCATTTGGTTCTCGGATCCAGGCA 680
Db      5228  CTTTGTACATCTCATTTGGTTCTCGGATCCAGGCA 5264

```

```

RESULT 12
US-09-867-701-8500
; Sequence 8500, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8500
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-8500

Query Match          20.5%; Score 234.6; DB 10; Length 339;
Best Local Similarity 98.0%; Pred. No. 4.4e-66;
Matches 248; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy      896  TTCACAGTTGATGCCCTTCTCTCTCTCTCAAAATCAATGTCCAGGAGTCAAAAGGGCTGT 955
          |||
Db       23  TTTCAACGTTGATGCCCTTCTCTTCTCTCAAAATCAATGTCCAGGAGTCAAAAGGGCTGT 82
          |||

Qy      956  AGCACAGATGGAGTTTGATTTATCCCTCTCTCCCCAACACACTAGGAATCGAATCTTTT 1015
          |||
Db       83  AGCACAGATGGAGTTTGATTTATCCCTCTCTCCCCAACACACTAGGAATCGAATCTTTT 142
          |||

```

QY	1016	CTTTTATTA-TTTTTTGAGATGGAGTCTGCTCTGTTGCCAGCTGAGTGCAGTGTGTGTA	1074
Db	143	CTTTTATTTTTTTGAGATGGAGTCTGCTCTGTTGCCAGCTGGAGTGAAGTGGTGTGA	202
QY	1075	TCTCAGCTTACTGCAACCTCTGTCTCCGGGTTCAAGCAAATCTCCCATCTCAGCCTCCT	1134
Db	203	TCTAAGCTTACTGCAACCTCTGTCTCAGGGTTCAAGCAAATCTCCCATCTCAGCCTCCT	262
QY	1135	GAGTAGCTGGGAT	1147
Db	263	GAGTAGCTGGAT	275

```

RESULT 13
US-09-814-353-4106/c
; Sequence 4106, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4106
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-4106
Query Match 12.0%; Score 137.6; DB 12; Length 510;
Best Local Similarity 96.6%; Pred. No. 3.3e-34;
Matches 140; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GCCTTGGGTCAAGCAGAAATTAATAGCGAGGGGAATGCACCTGTAGTGGGCGCTA 60
Db 163 GCCTTGGGTCAAGCAGAAATTAATAGCGAGGGGAATGCACCTGTAGTGGGCGCTA 104

QY 61 CTGCGACGCTGAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGTTCTCTGAGGATGA 120
Db 103 CTGCGACGCTGAAGAGAGAGTGGCTCGACTCTTGACCATGGCTGGGTTCTCTGAGGATGA 44

QY 121 GCTAAACCCCTTCCATGTACTGGGG 145
Db 43 GCTAAACCCCTTCCATGTACTGGGG 19

RESULT 14
US-09-814-353-10412/c
; Sequence 10412, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:

```

```
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10412
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-10412
```

```
Query Match 12.0%; Score 137.6; DB 12; Length 510;
Best Local Similarity 96.6%; Pred. No. 3.3e-34;
Matches 140; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCCTTGGGTCAAGCAGAGATATTATAGGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 60
Db 163 GCCTTGGGTCAAGCAGAGATATTATAGGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 104

Qy 61 CTGCCAGCCTGAAGAGGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCTGAGGATGA 120
Db 103 CTGCCAGCCTGGAGAGGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCTGAGGATGA 44

Qy 121 GCTAAACCCCTTCCATGTACTGGGG 145
Db 43 GCTAAACCCCTTCCATGTACTGGG 19
```

```
RESULT 15
US-09-814-353-16796/c
; Sequence 16796, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
```

```
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16796
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-16796

Query Match 11.9%; Score 137; DB 12; Length 598;
Best Local Similarity 96.6%; Pred. No. 5.7e-34;
Matches 140; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GCCTTGGGTCAAGCAGAGATATTATAGGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 60
Db 234 GCCTTGGGTCAAGCAGAGATATTATAGGCAGGGGAATGCACCTGTAGCTAGTGGCGCTA 175

Qy 61 CTGCCAGCCTGAAGAGGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCTGAGGATGA 120
Db 174 CTGCCAGCCTGGAGAGGAGTGGCTCGACTTTGACCATGGCTGGGGTTCTCTGAGGATGA 115

Qy 121 GCTAAACCCCTTCCATGTACTGGGG 145
Db 114 GCTAAACCCCTTACCATGTACTCGG 90

Search completed: December 1, 2003, 10:09:53
Job time : 1870 secs
```